# The cumulative cultural evolution of category structure in an open-ended meaning space

**Jon W. Carr**
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

**Hannah Cornish**
Department of Psychology
University of Stirling

**Simon Kirby**
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

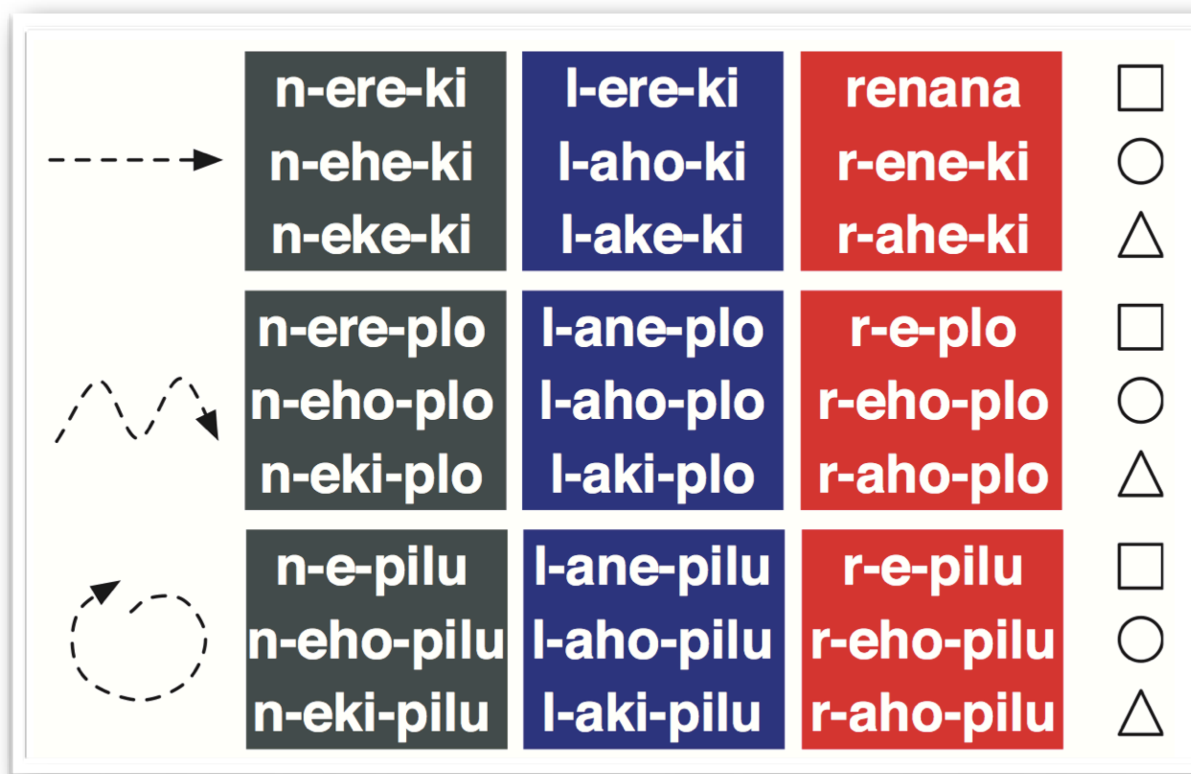
# Recap of Kirby et al. (2008)

Showed that the cultural transmission of language can give rise to the same structural properties we find in natural languages

The meanings form a 3 × 3 × 3 space in which each of three dimensions vary over three discrete categories

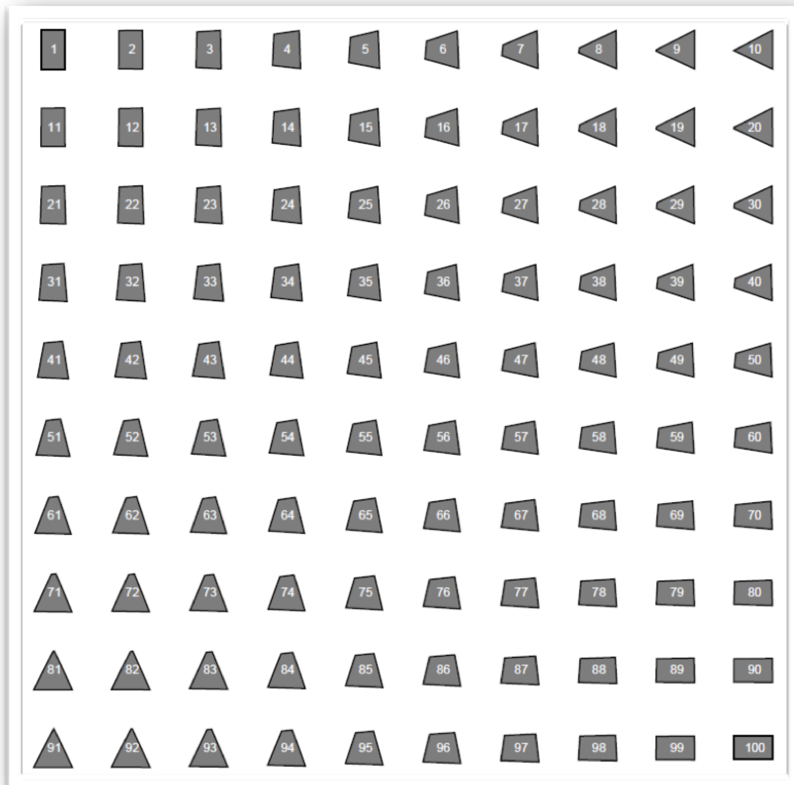
But this is not a realistic representation of the real world

The human conception of the world is higher-dimensional, continuous, and open-ended.

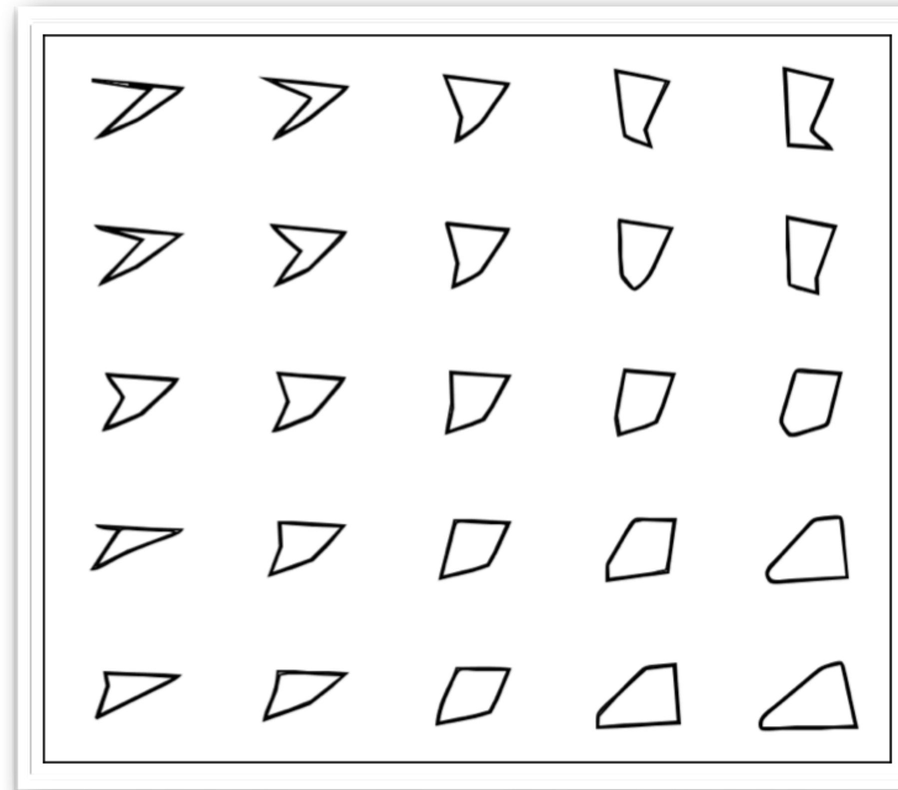


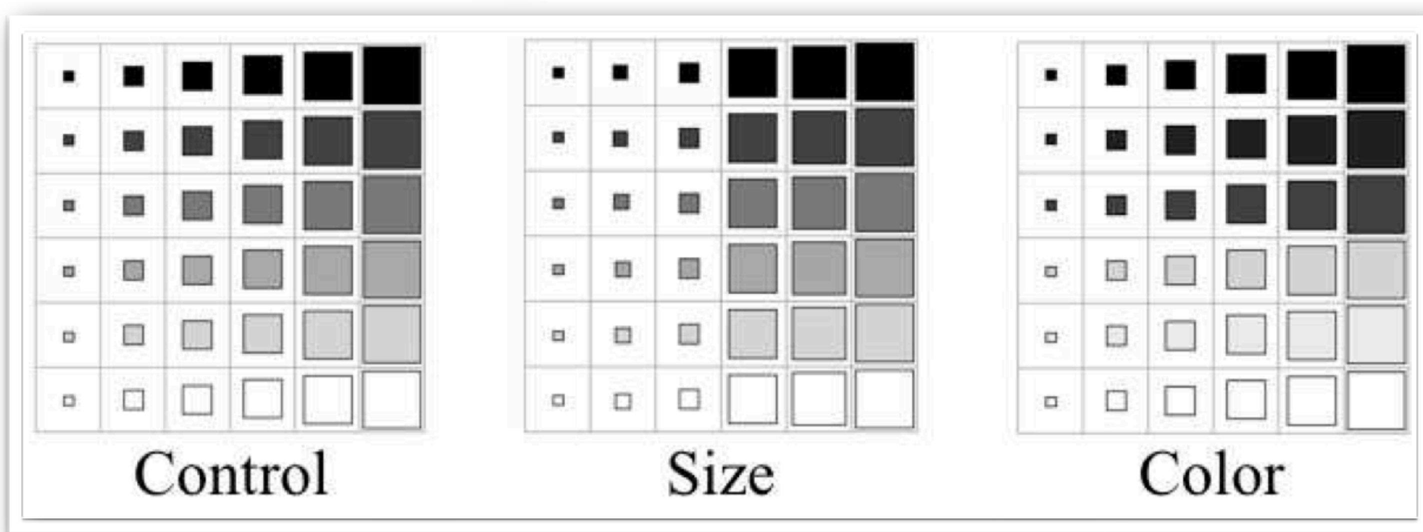
# Continuous spaces in previous work

Matthews (2009)



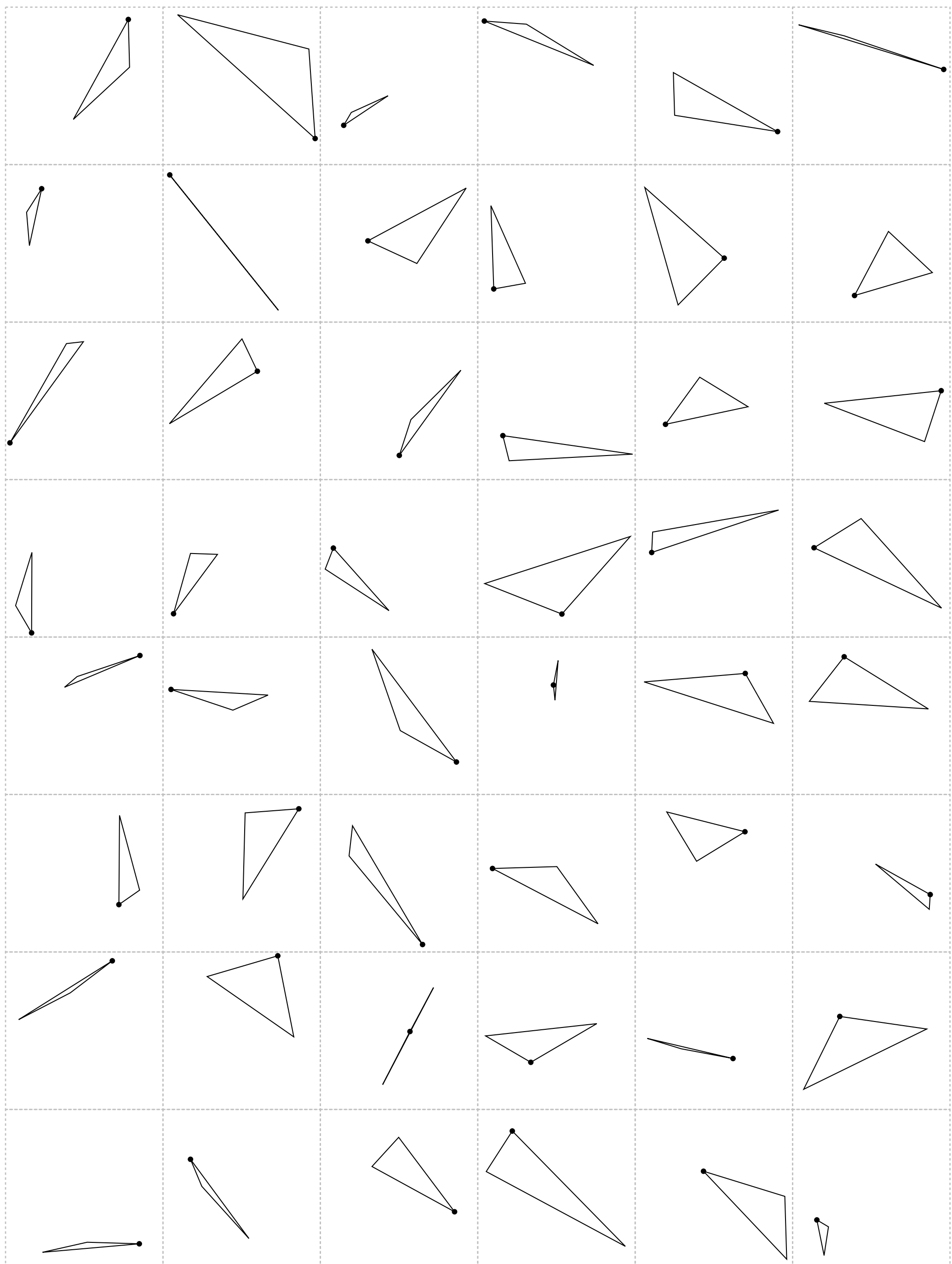
Silvey, Kirby, & Smith (2013)

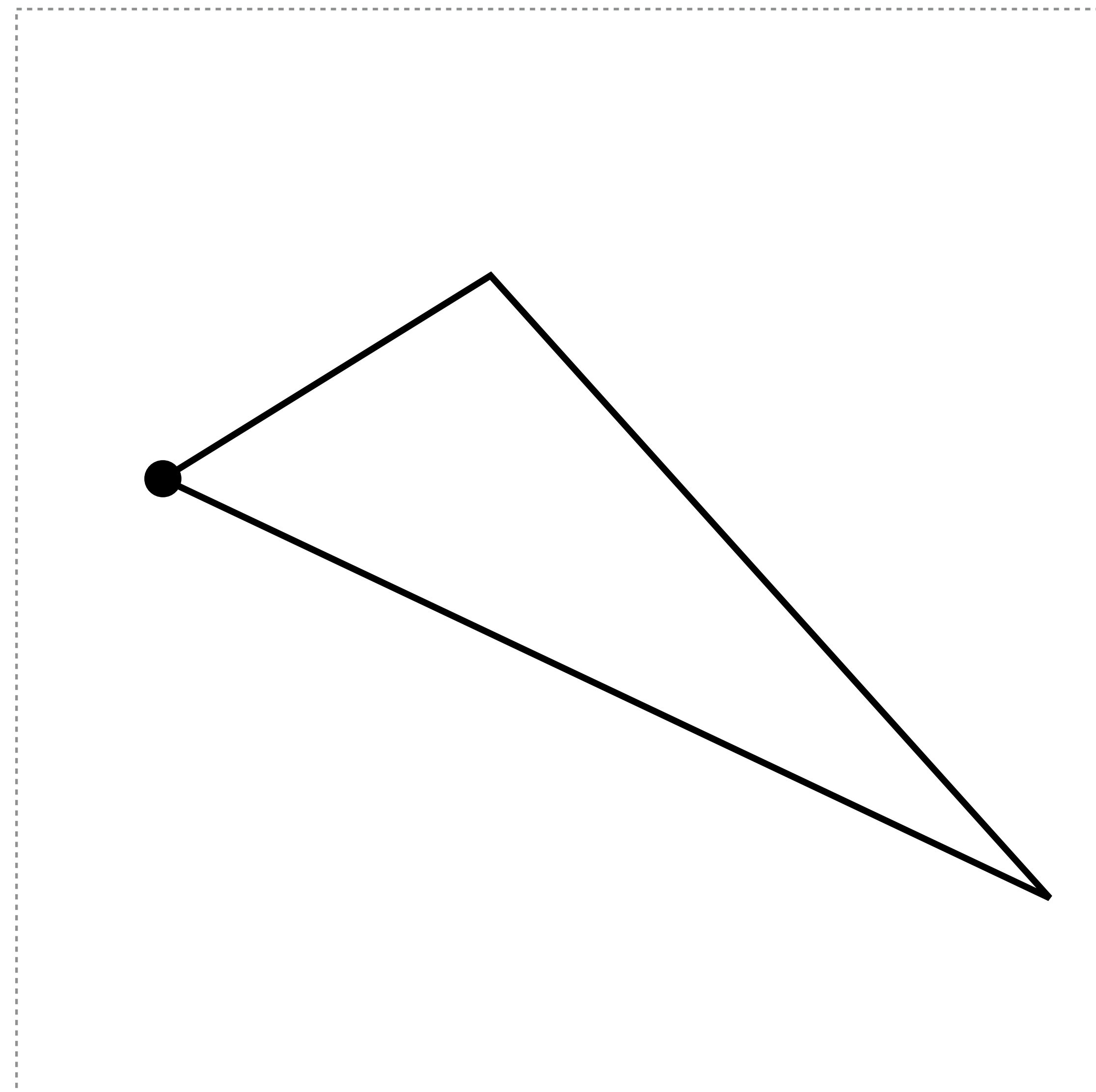





Perfors & Navarro (2011)

# Triangle stimuli





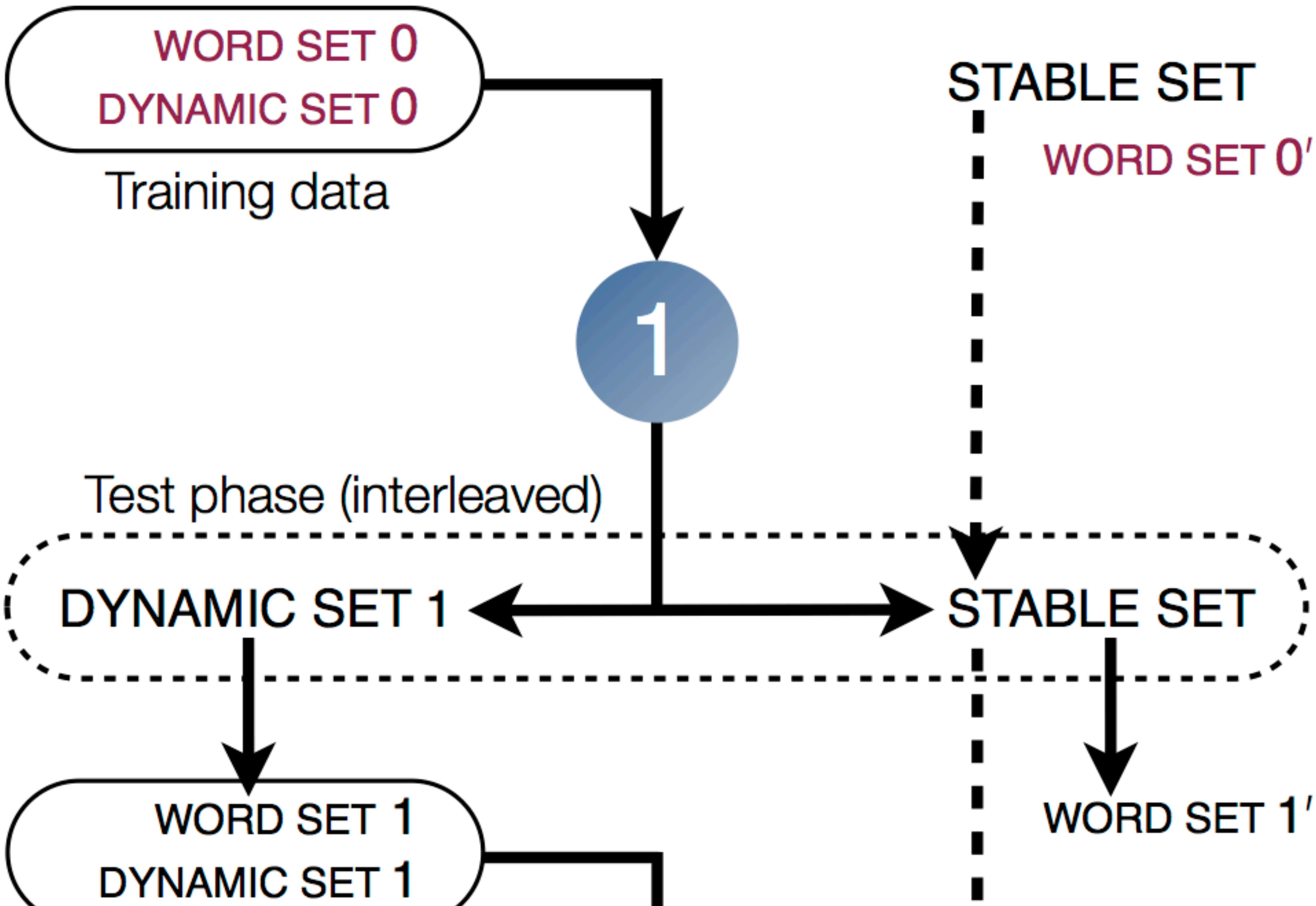
# Linguistic stimuli

Initial word sets generated randomly from the set of consonants {d, f, k, m, p, z} and the set of vowels {a, i, o, u}

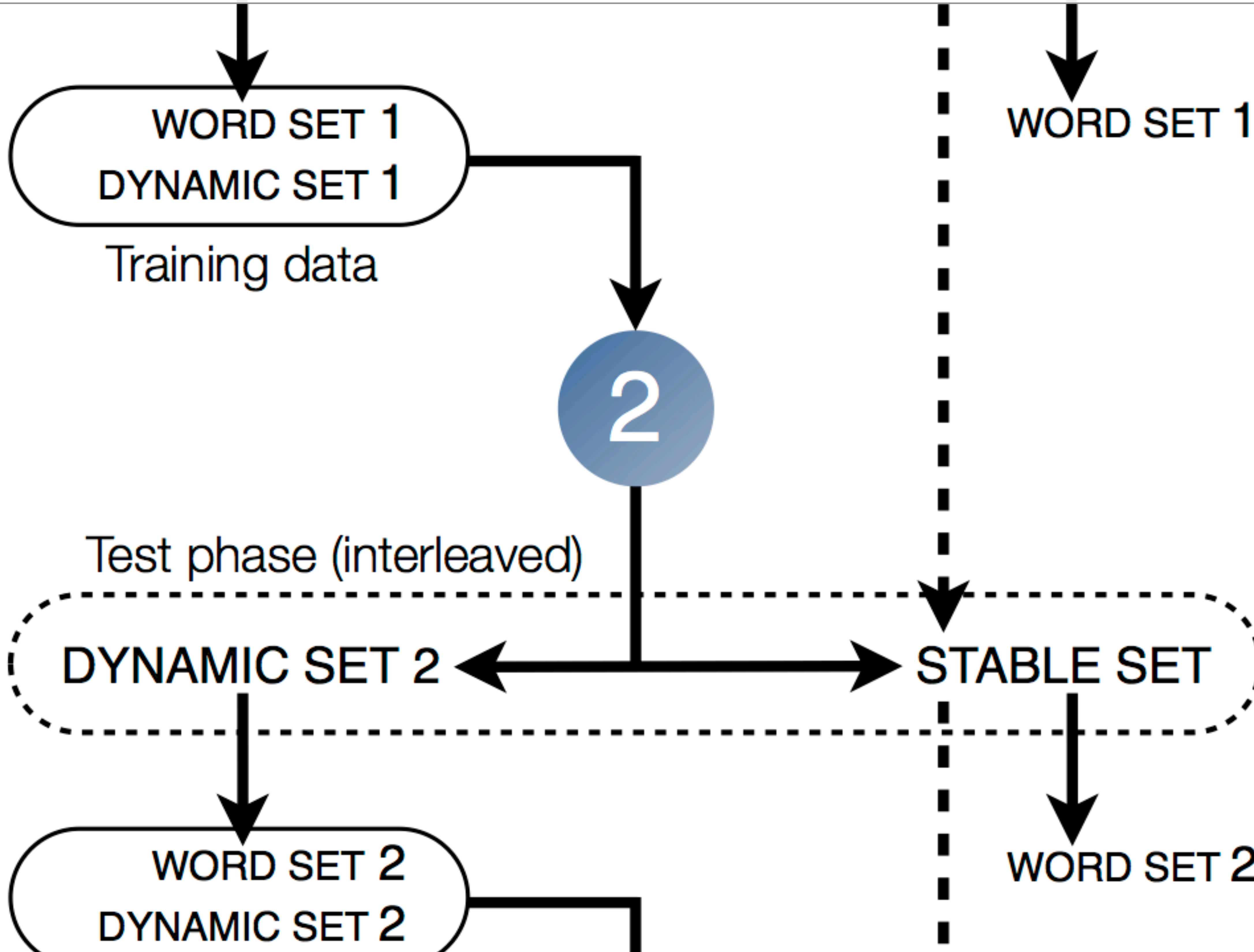
Words consisted of between 2 and 4 syllables

The presentation of the words was accompanied by a vocal rendition produced with a speech synthesizer

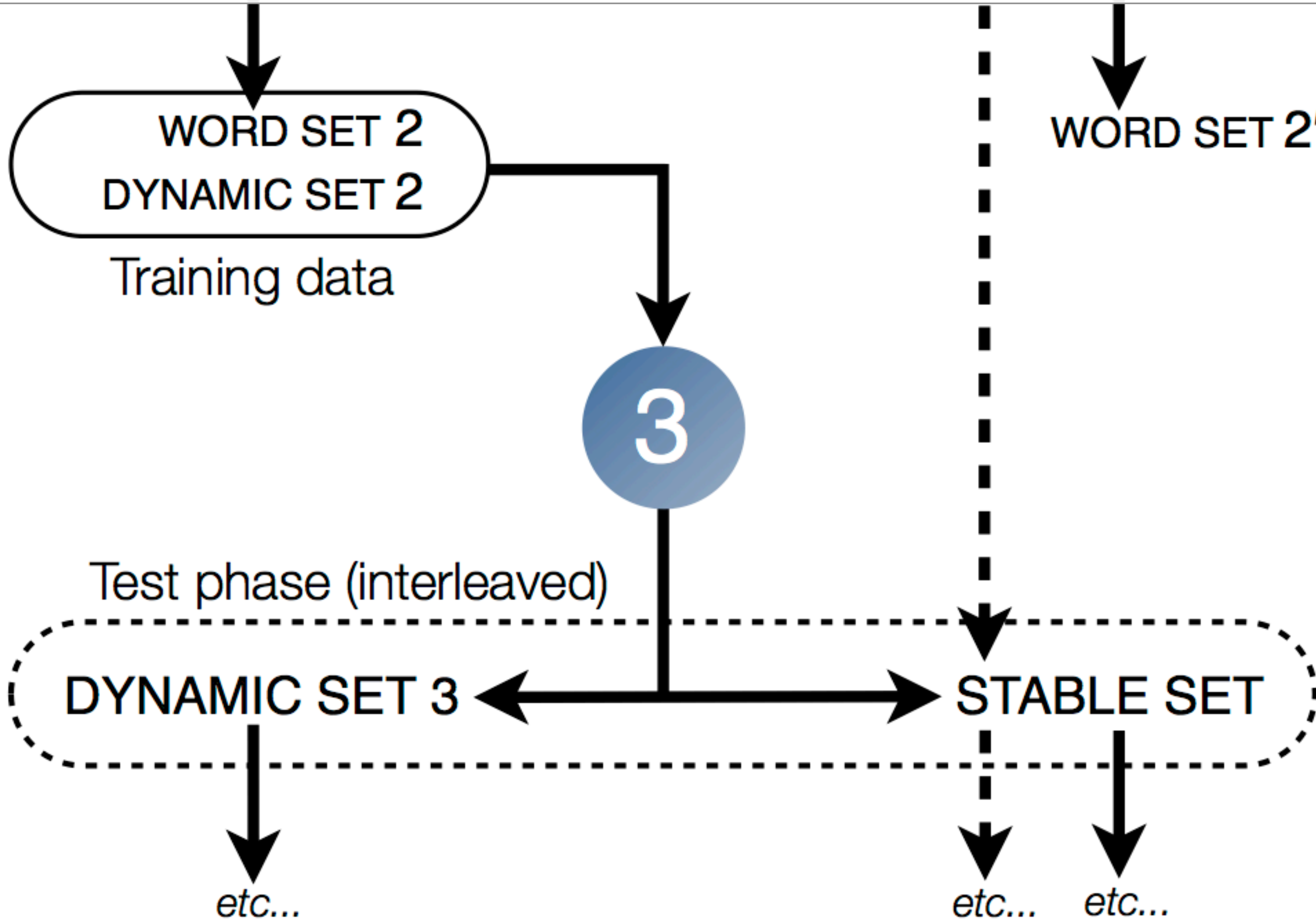
# Procedure


WORD SET 0
DYNAMIC SET 0
Training data
STABLE SET
WORD SET 0′
1
Test phase (interleaved)
DYNAMIC SET 1
STABLE SET
WORD SET 1
DYNAMIC SET 1
WORD SET 1′

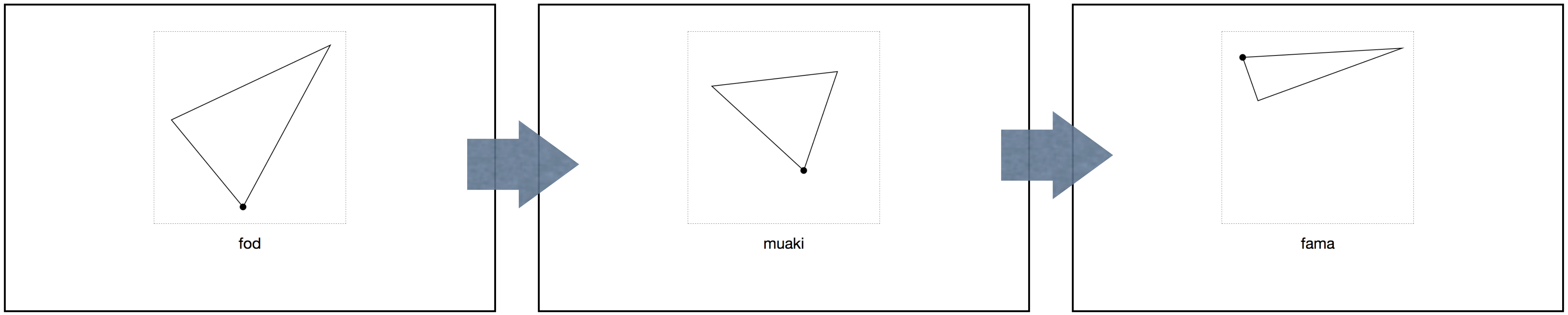
## Procedure


WORD SET 1
DYNAMIC SET 1
Training data
WORD SET 1′
2
Test phase (interleaved)
DYNAMIC SET 2
STABLE SET
WORD SET 2
DYNAMIC SET 2
WORD SET 2′

# Procedure

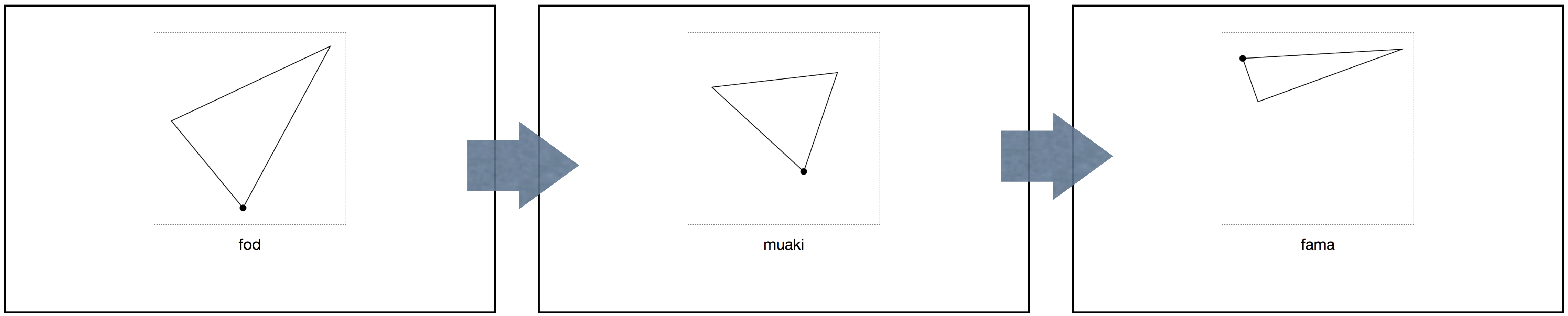

WORD SET 2
DYNAMIC SET 2
Training data
WORD SET 2′
3
Test phase (interleaved)
DYNAMIC SET 3
STABLE SET
etc...
etc...
etc...

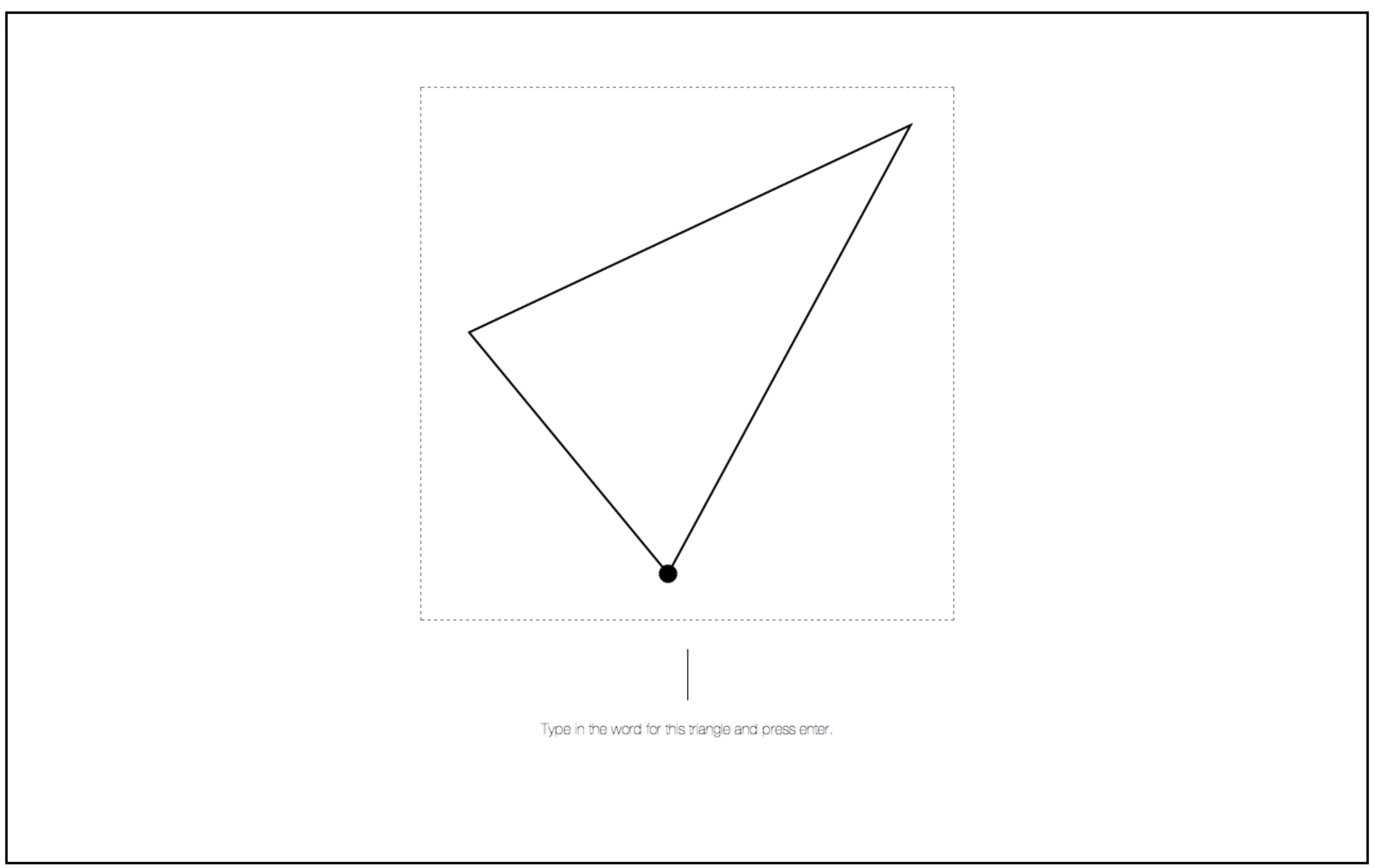
# Experiment interface: Training




Three stimuli presented from the dynamic set for 5 seconds each

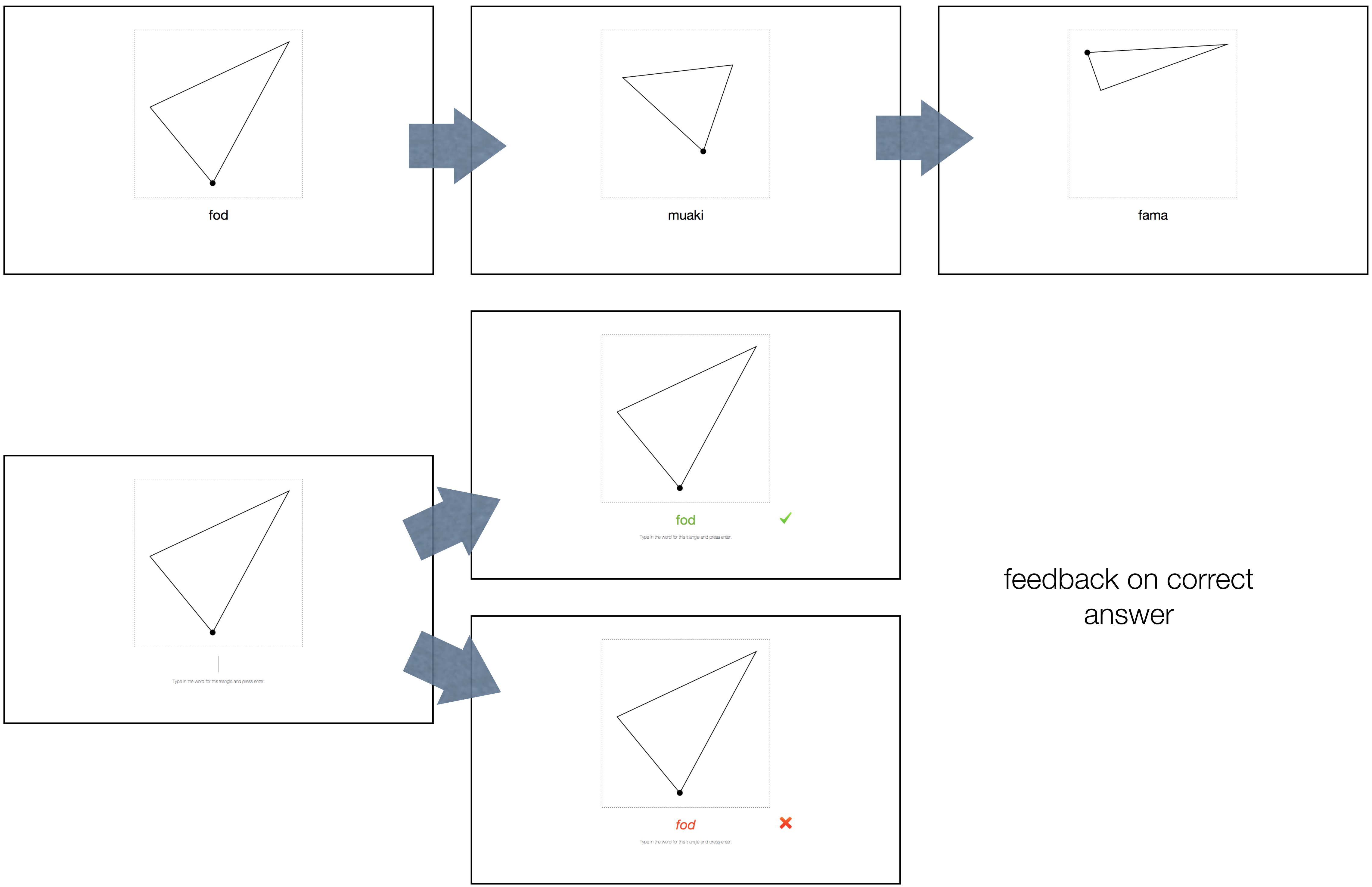
# Experiment interface: Training



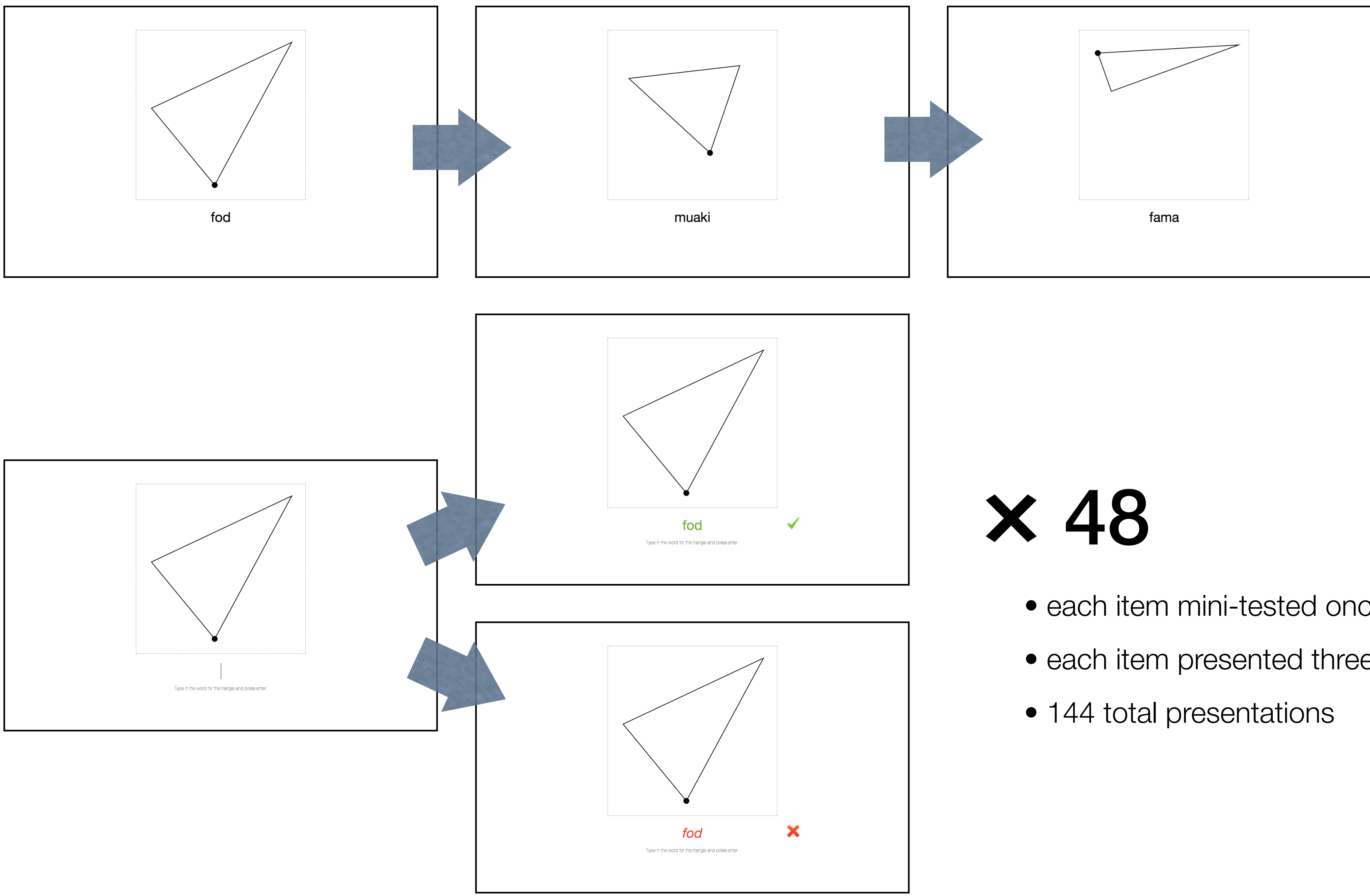

"mini test" on one of the previous three stimuli

# Experiment interface: Training




feedback on correct answer

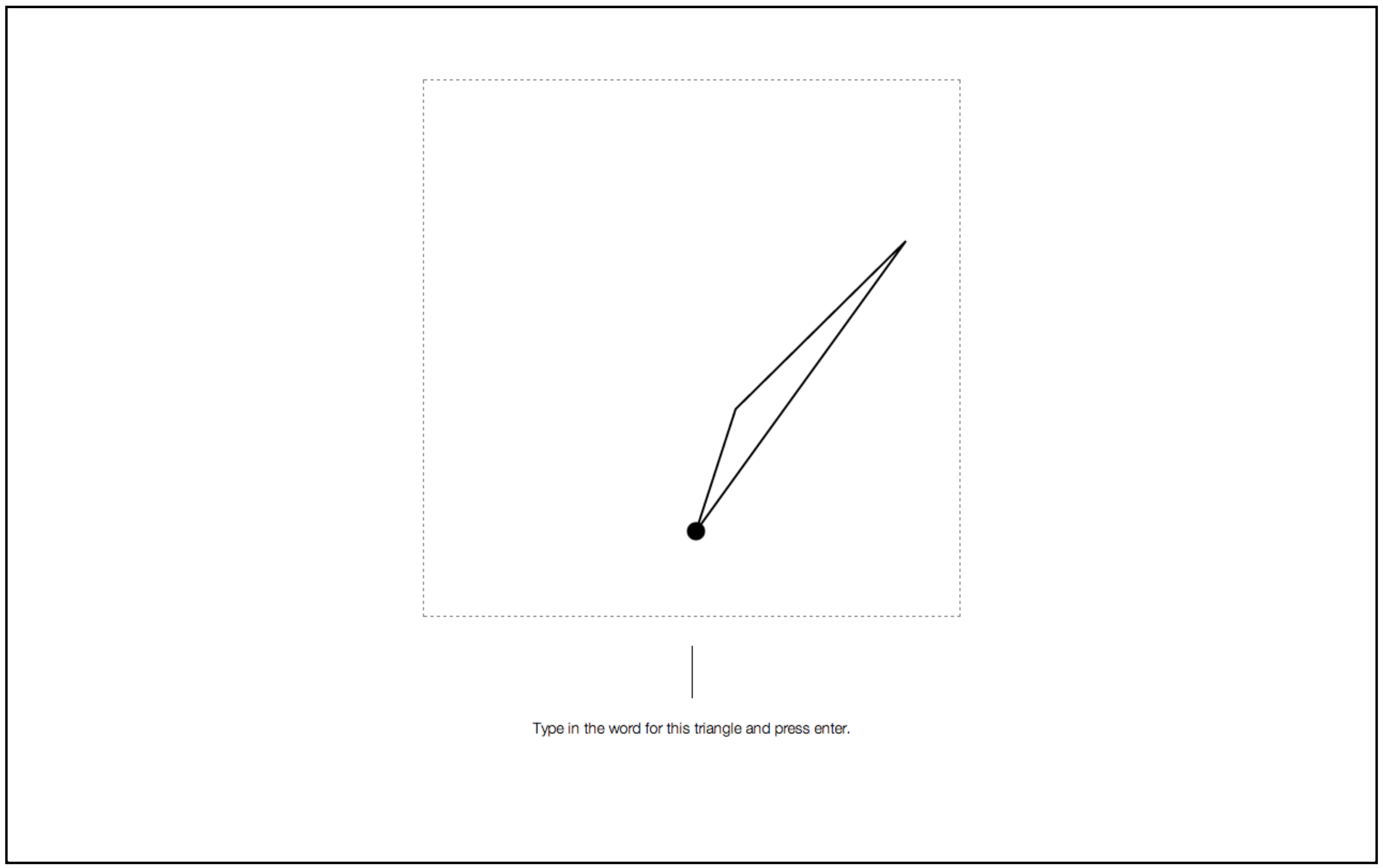
# Experiment interface: Training




**× 48**

- each item mini-tested once
- each item presented three times
- 144 total presentations

# Experiment interface: Testing




**× 96**

- 48 items from stable set
- 48 items from dynamic set
- interleaved

# Measure of learnability

Transmission error is used as a proxy for learnability

Measured only on the stable set of items for consistency across generations

Greater error in predicting the words that the previous participant applied to items in the stable set implies a less learnable language (and vice versa)

Transmission error is the mean normalized Levenshtein distance:

$$E(i) = \frac{1}{|M|} \sum_{m \in M} \frac{\mathrm{LD}(s_i^m, s_{i-1}^m)}{\max(\mathrm{len}(s_i^m), \mathrm{len}(s_{i-1}^m))}$$

# Measure of structure

The languages are essentially mappings between signals and meanings

To measure structure, we correlate the dissimilarity between pairs of strings with the dissimilarity between pairs of triangles for all $n(n-1)/2$ pairs

We then perform a Mantel test (Mantel, 1967) which compares this correlation against a distribution of correlations for 50,000 Monte-Carlo permutations of the signal-meaning pairs

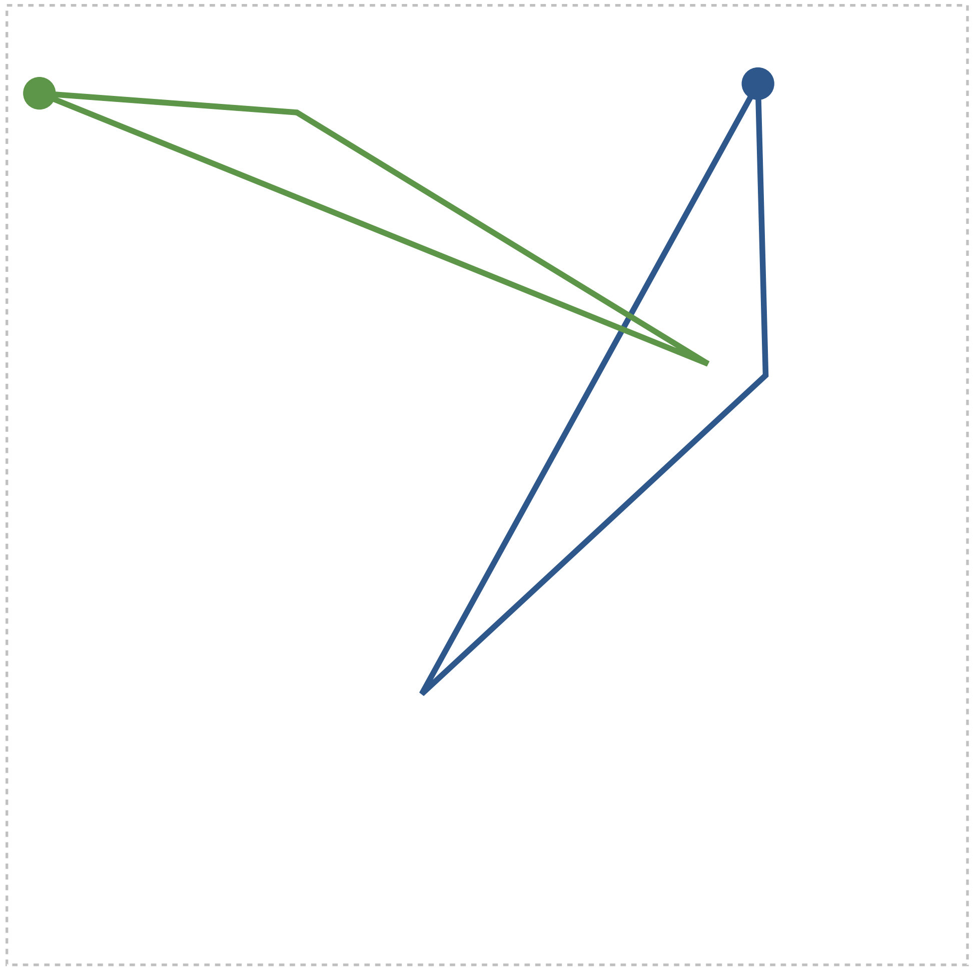
This yields a standard score ($z$-score) quantifying the significance of the observed correlation

Normalized Levenshtein distance used to measure the dissimilarity between pairs of strings

# Triangle dissimilarity metric

The dissimilarity between two triangles is taken as the sum of Euclidean distances between vertices

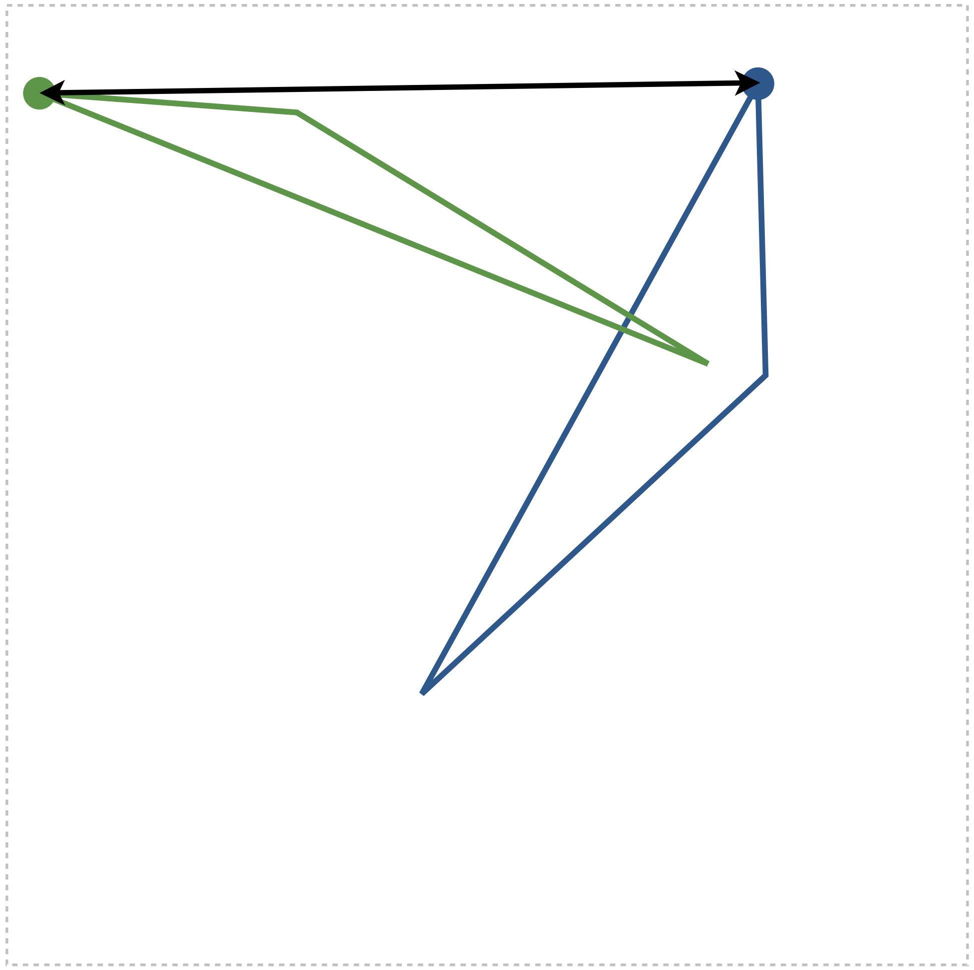
$$d_T(A, B) = d_E(A_1, B_1) + \min[d_E(A_2, B_2) + d_E(A_3, B_3), d_E(A_2, B_3) + d_E(A_3, B_2)]$$



# Triangle dissimilarity metric

The dissimilarity between two triangles is taken as the sum of Euclidean distances between vertices

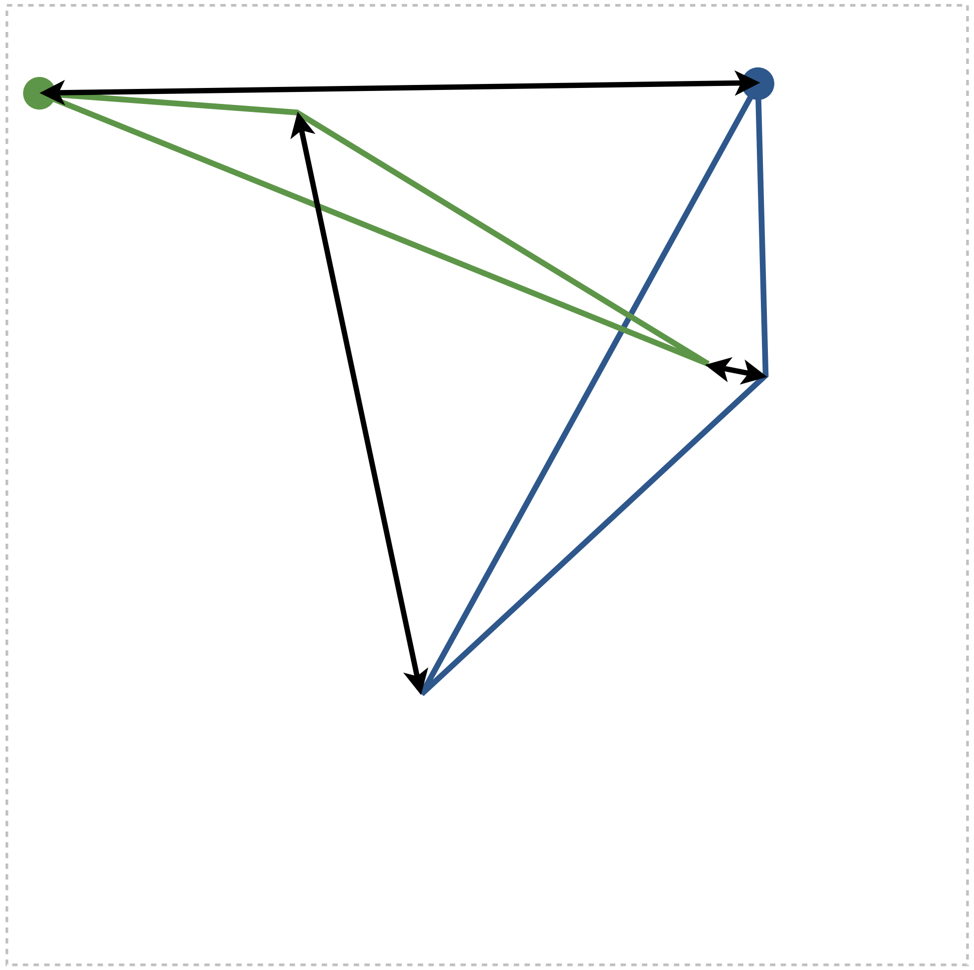
$$d_T(A,B) = d_E(A_1,B_1) + \min[d_E(A_2,B_2) + d_E(A_3,B_3), d_E(A_2,B_3) + d_E(A_3,B_2)]$$



# Triangle dissimilarity metric

The dissimilarity between two triangles is taken as the sum of Euclidean distances between vertices

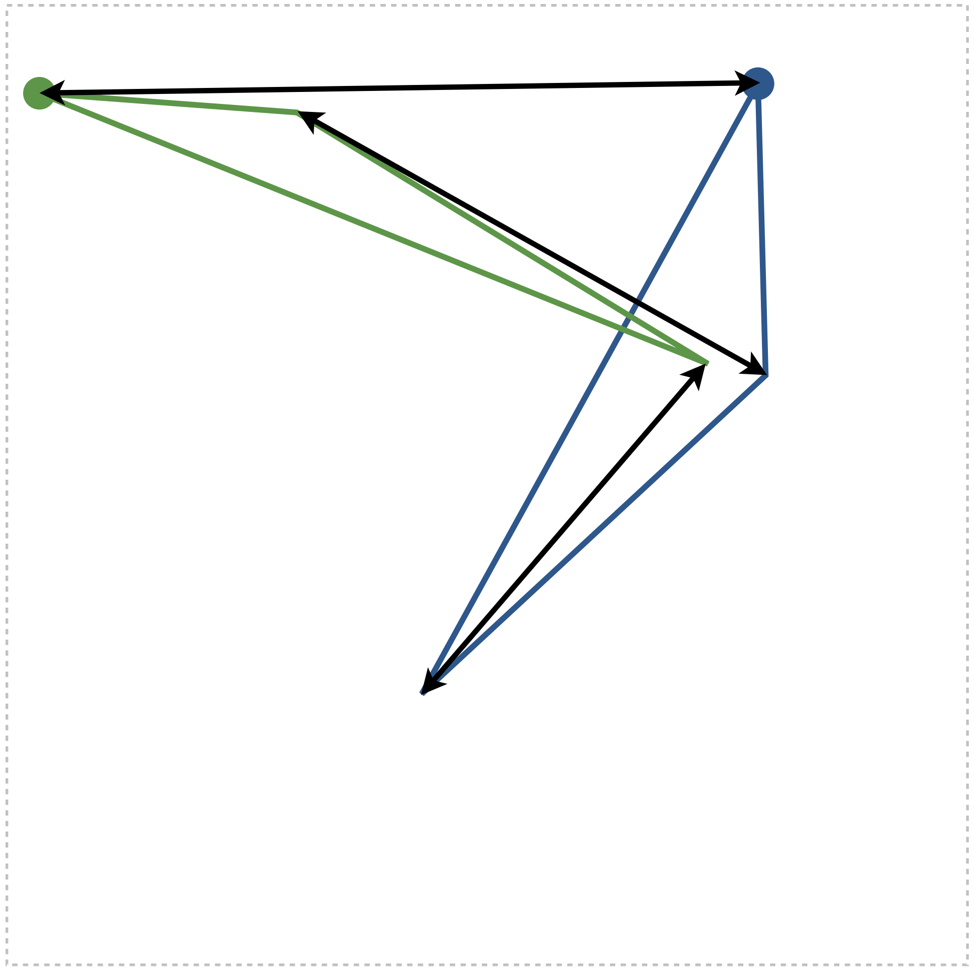
$$d_T(A,B) = d_E(A_1,B_1) + \min[d_E(A_2,B_2) + d_E(A_3,B_3), d_E(A_2,B_3) + d_E(A_3,B_2)]$$



# Triangle dissimilarity metric

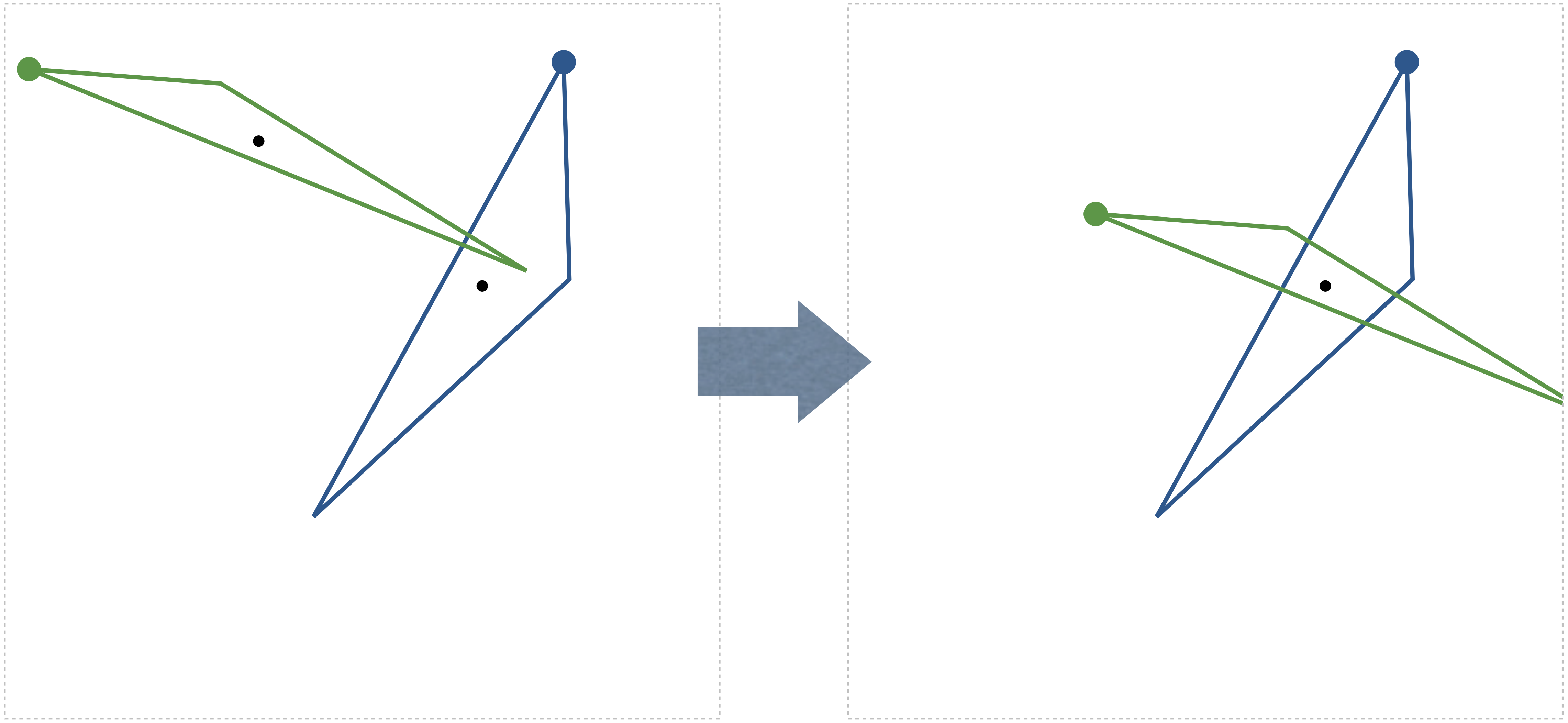
The dissimilarity between two triangles is taken as the sum of Euclidean distances between vertices

$$d_T(A,B) = d_E(A_1,B_1) + \min[d_E(A_2,B_2) + d_E(A_3,B_3), d_E(A_2,B_3) + d_E(A_3,B_2)]$$



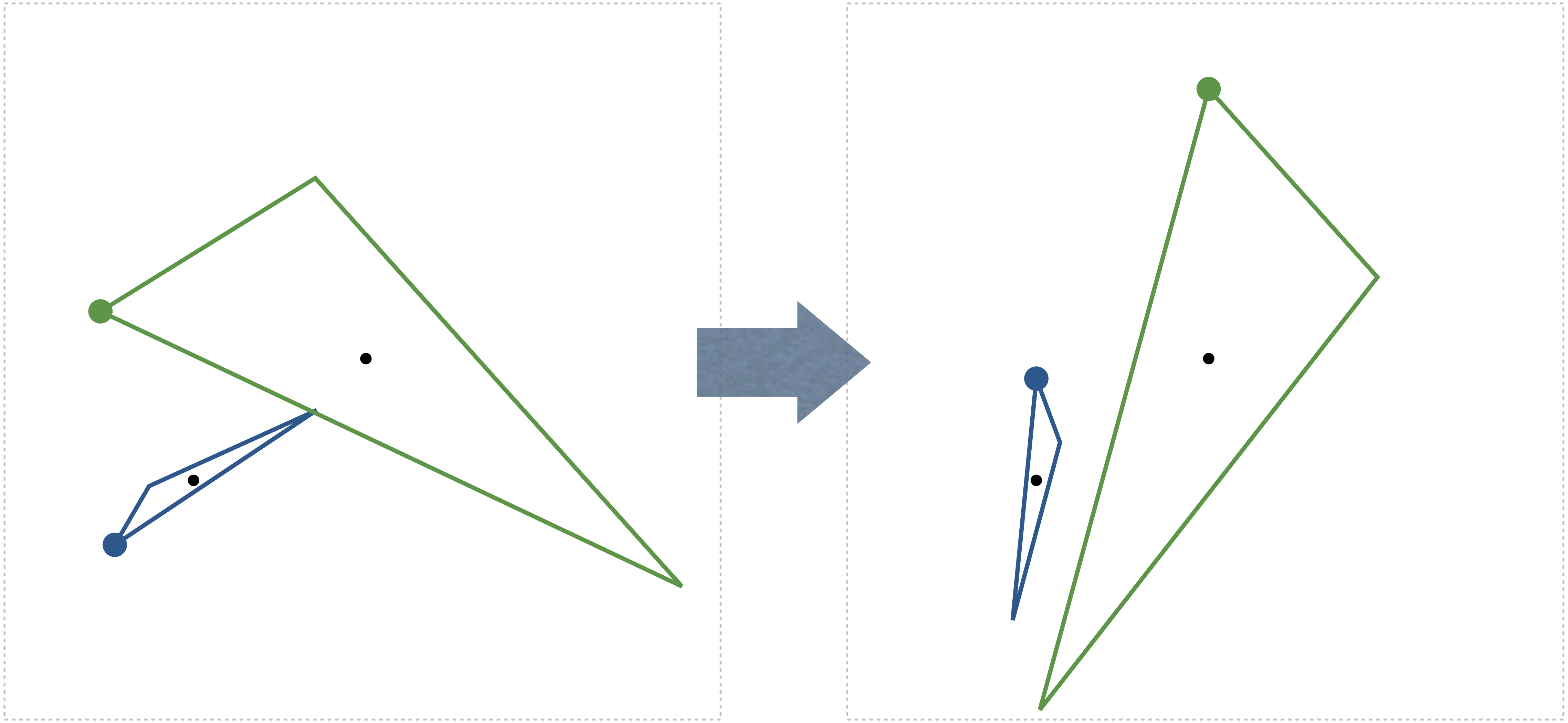
# Triangle dissimilarity metric

***$d_T$* up to translation:** The triangles are translated to the same location in the plane based on their centroids



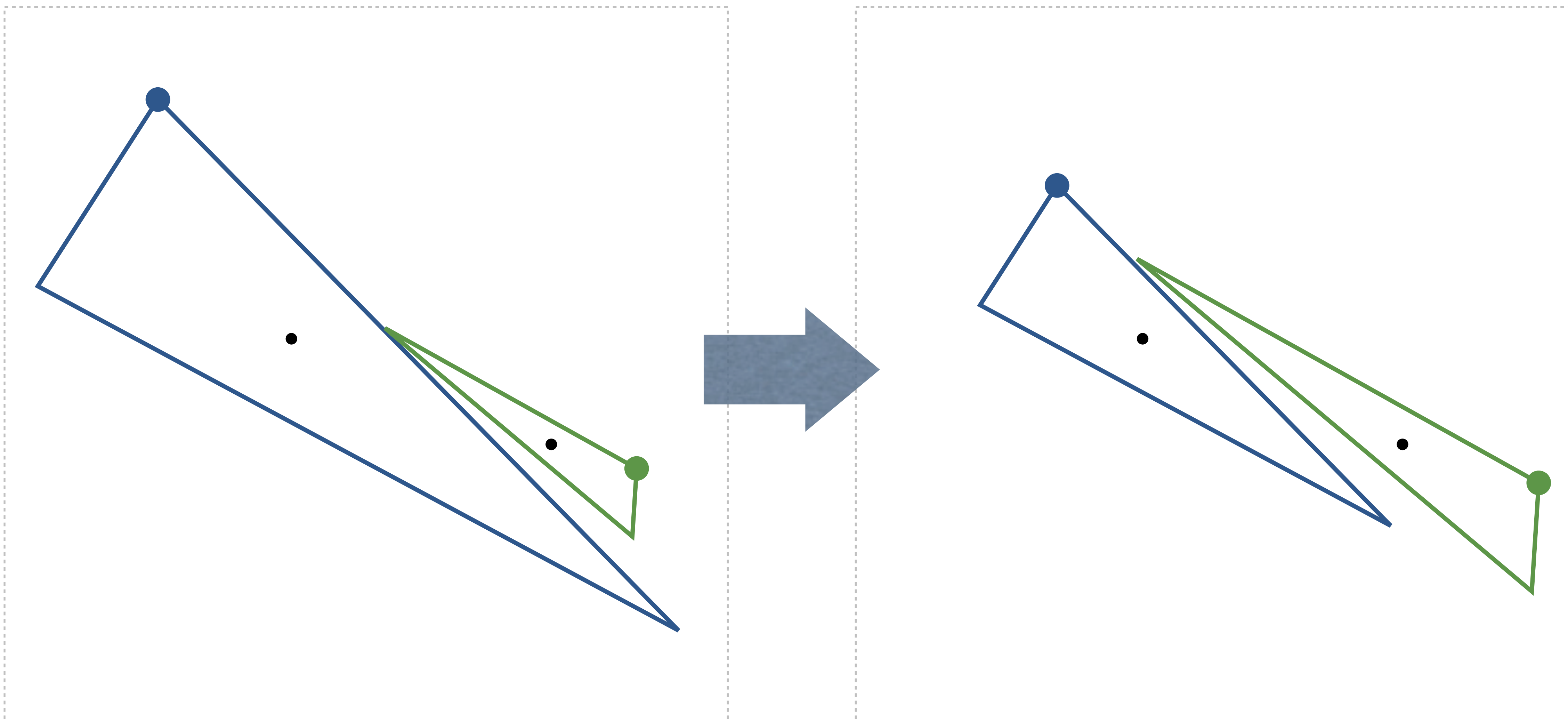
# Triangle dissimilarity metric

***$d_T$* up to rotation:** The triangles are rotated around their centroids so that they both “point” upwards



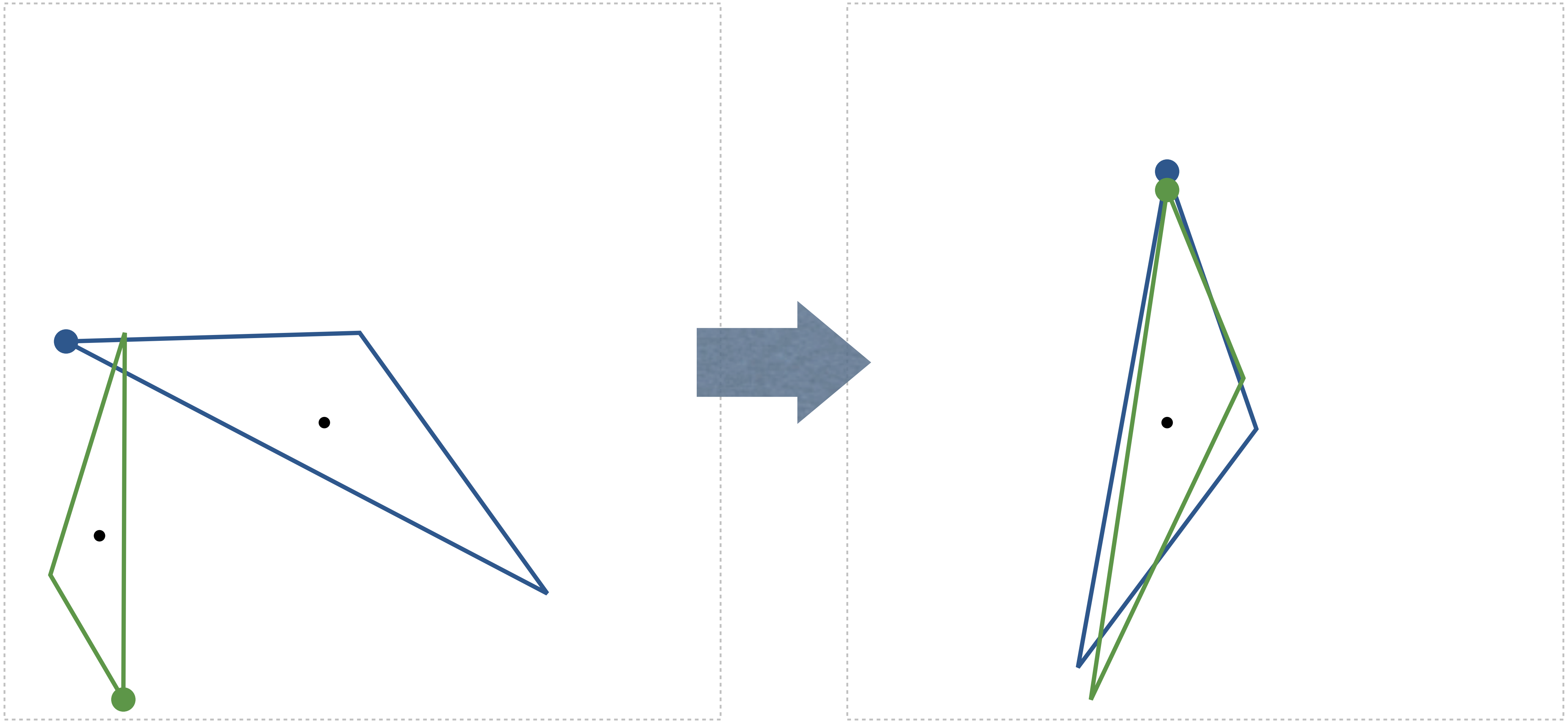
# Triangle dissimilarity metric

**$d_T$ up to scale:** The triangles are scaled around their centroids so that they have equal perimeter



# Triangle dissimilarity metric

***d*$_T$ up to scaled rigid motion:** The triangles are translated to the same location, rotated to the same direction, and scaled to the same size



# Triangle dissimilarity metric

List of eight triangle distance metrics alongside the geometrical properties that they ignore and consider

| Distance metric | Properties ignored | Properties considered |
|---|---|---|
| $d_T$ | — | shape, location, orientation, size |
| $d_T$ up to translation | location | shape, orientation, size |
| $d_T$ up to rotation | orientation | shape, location, size |
| $d_T$ up to scale | size | shape, location, orientation |
| $d_T$ up to rigid motion | location, orientation | shape, size |
| $d_T$ up to scaled translation | location, size | shape, orientation |
| $d_T$ up to scaled rotation | orientation, size | shape, location |
| $d_T$ up to scaled rigid motion | location, orientation, size | shape |

# Hypotheses

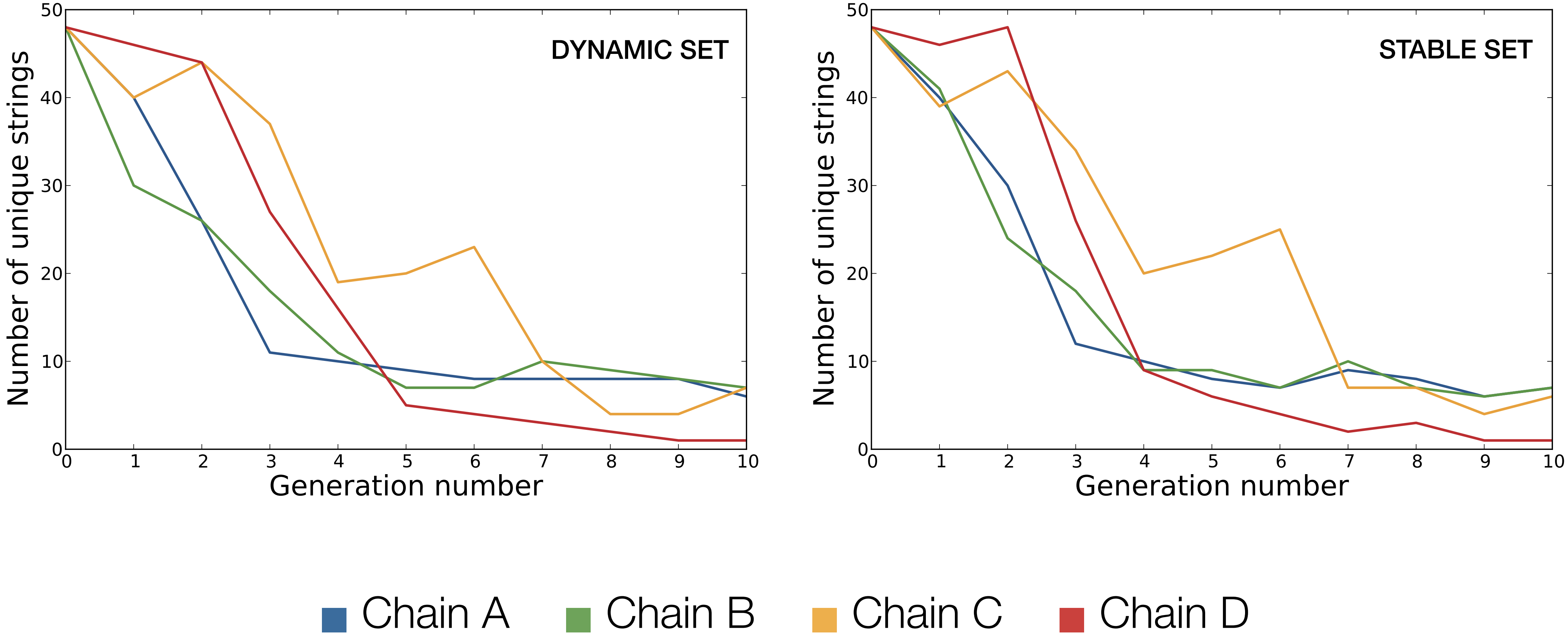
**Hypothesis 1:** the languages will become increasingly learnable over the course of the cultural generations

**Hypothesis 2:** categorical structure will emerge as a mechanism for circumventing the bottleneck on transmission

**Hypothesis 3:** given that Hypothesis 1 and Hypothesis 2 are supported, an increase in learnability will be explained by an increase in structure

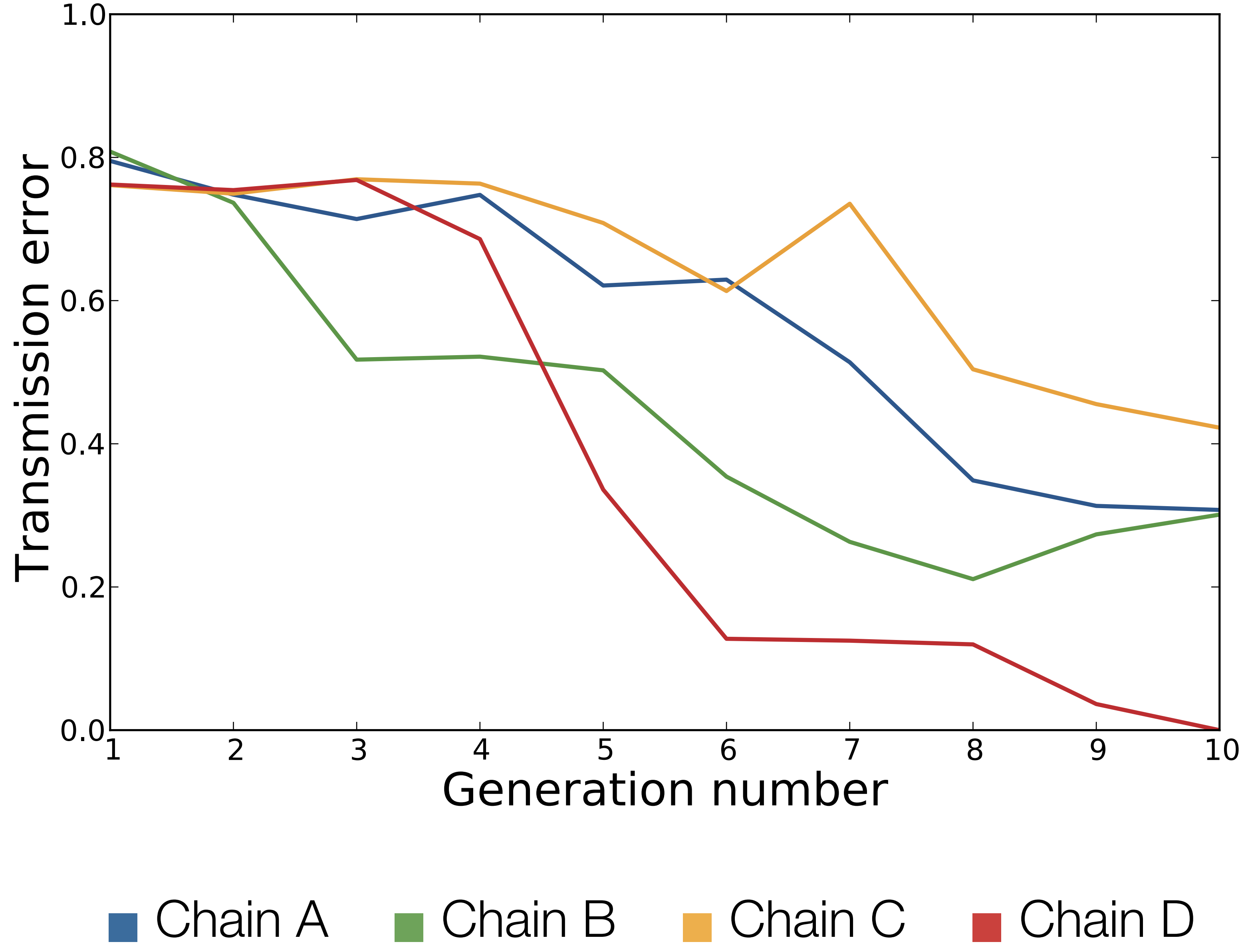
# Results: Unique strings

The number of unique strings in the dynamic and stable sets over the 10 generations for each chain

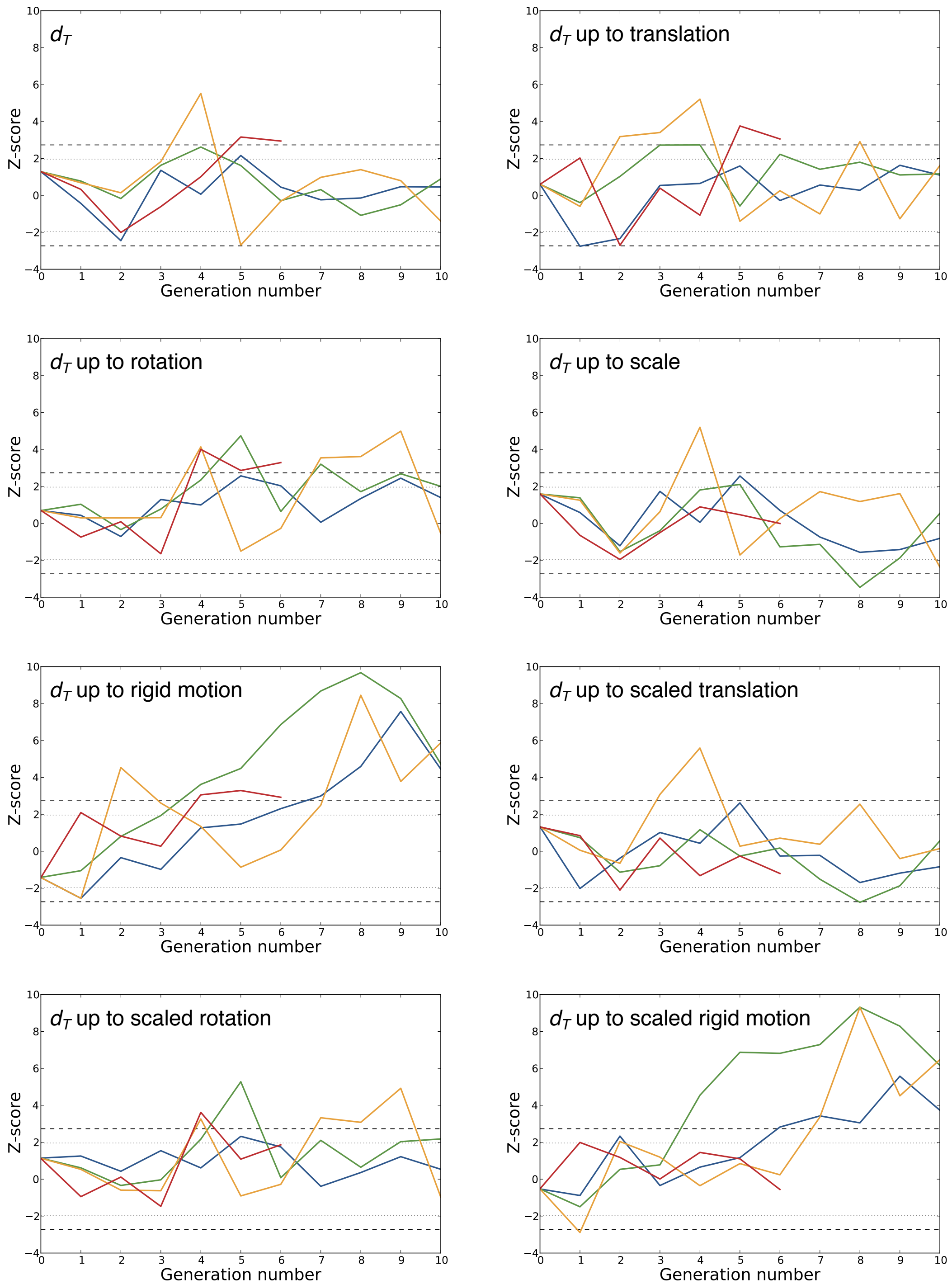


# Results: Learnability

Transmission error over 10 generations for each chain



# Results: Structure

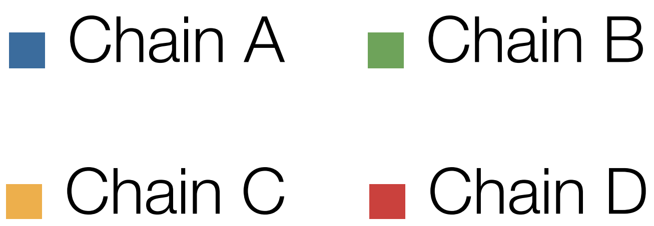



Structure results for the eight triangle dissimilarity metrics

Two metrics stand out in particular

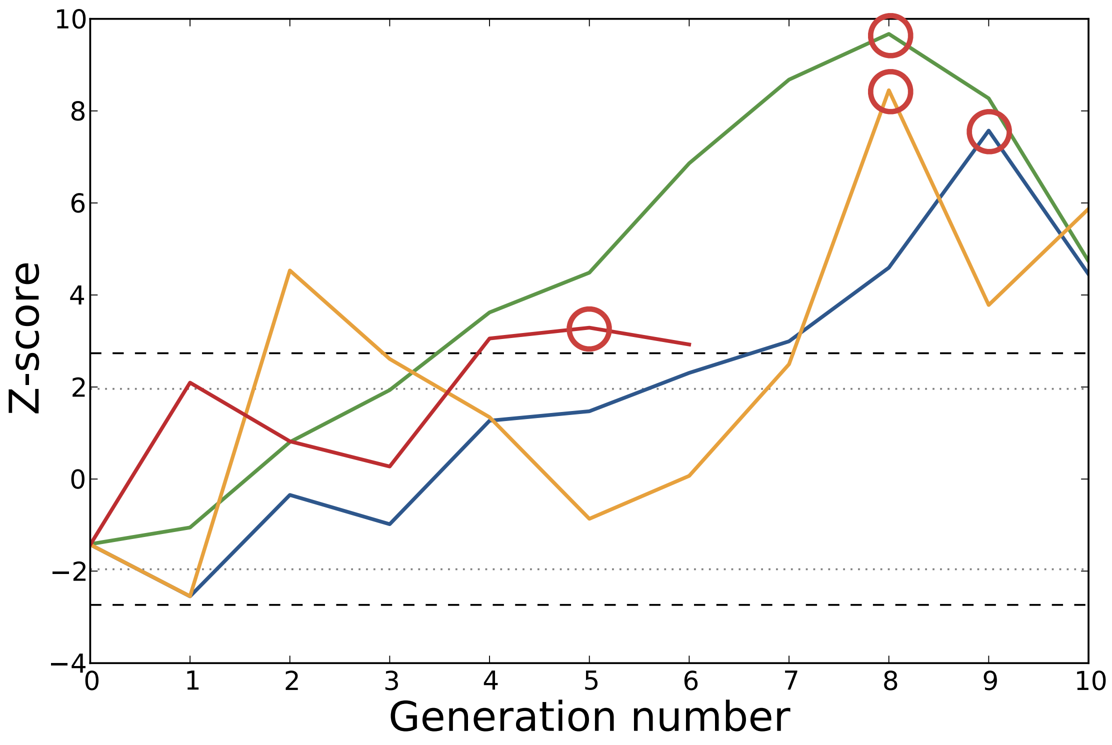
- $d_T$ up to rigid motion
- $d_T$ up to scaled rigid motion

These are the metrics that consider shape and size

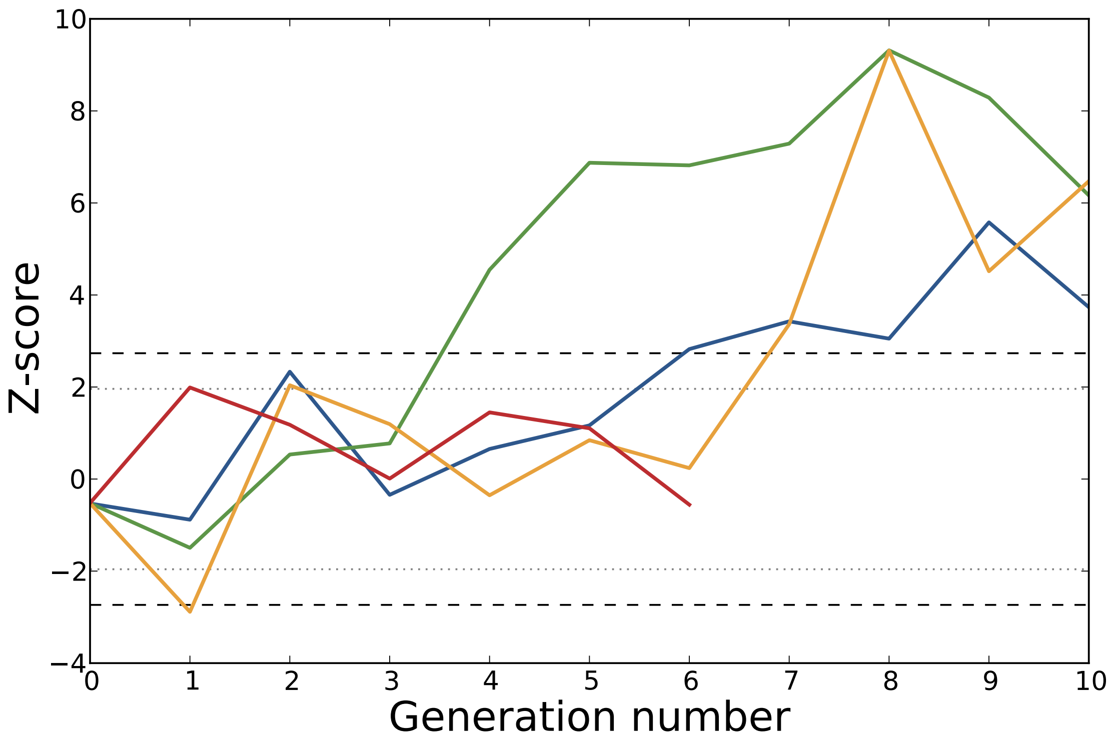


# Results: Structure

$d_T$ up to rigid motion

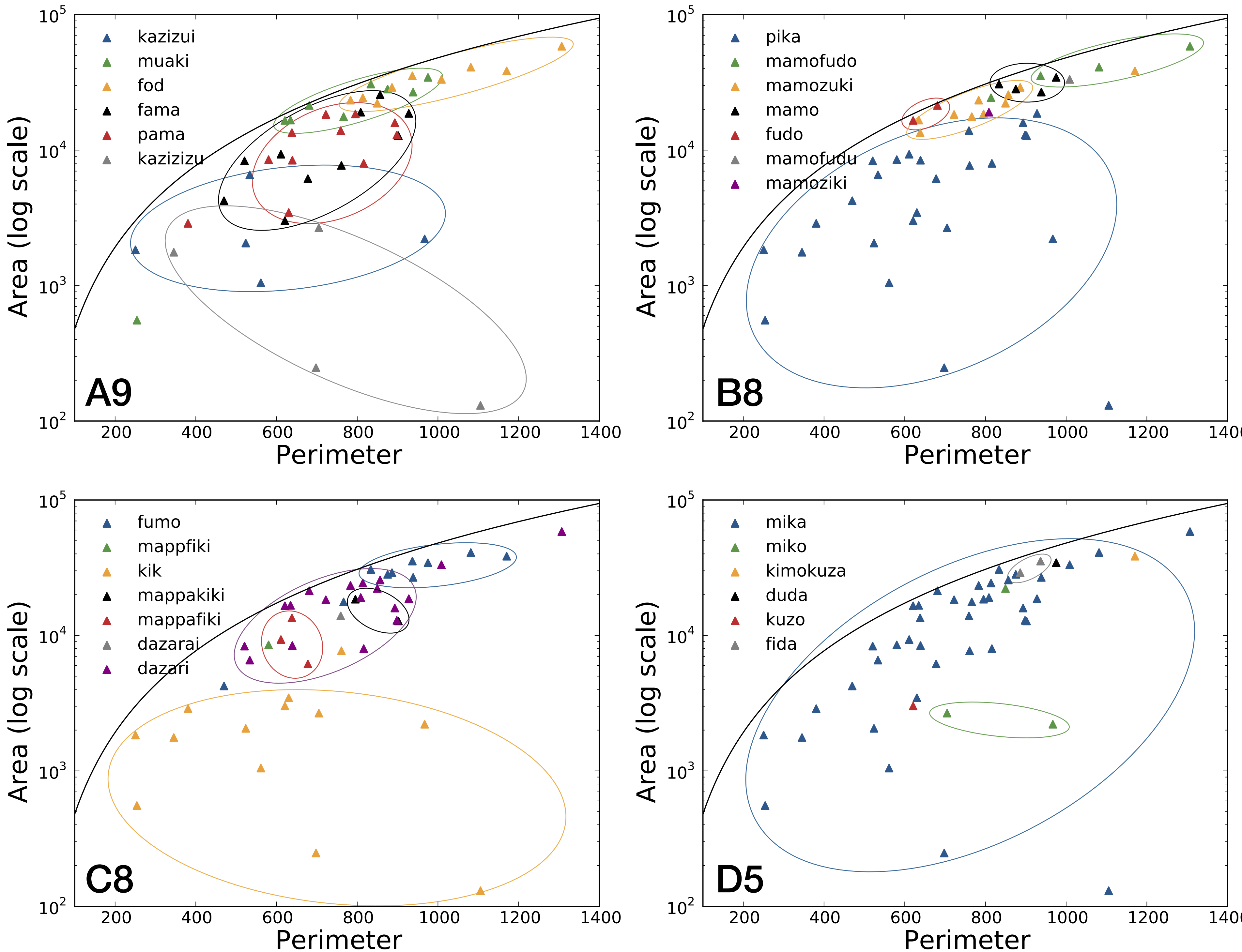



$d_T$ up to scaled rigid motion

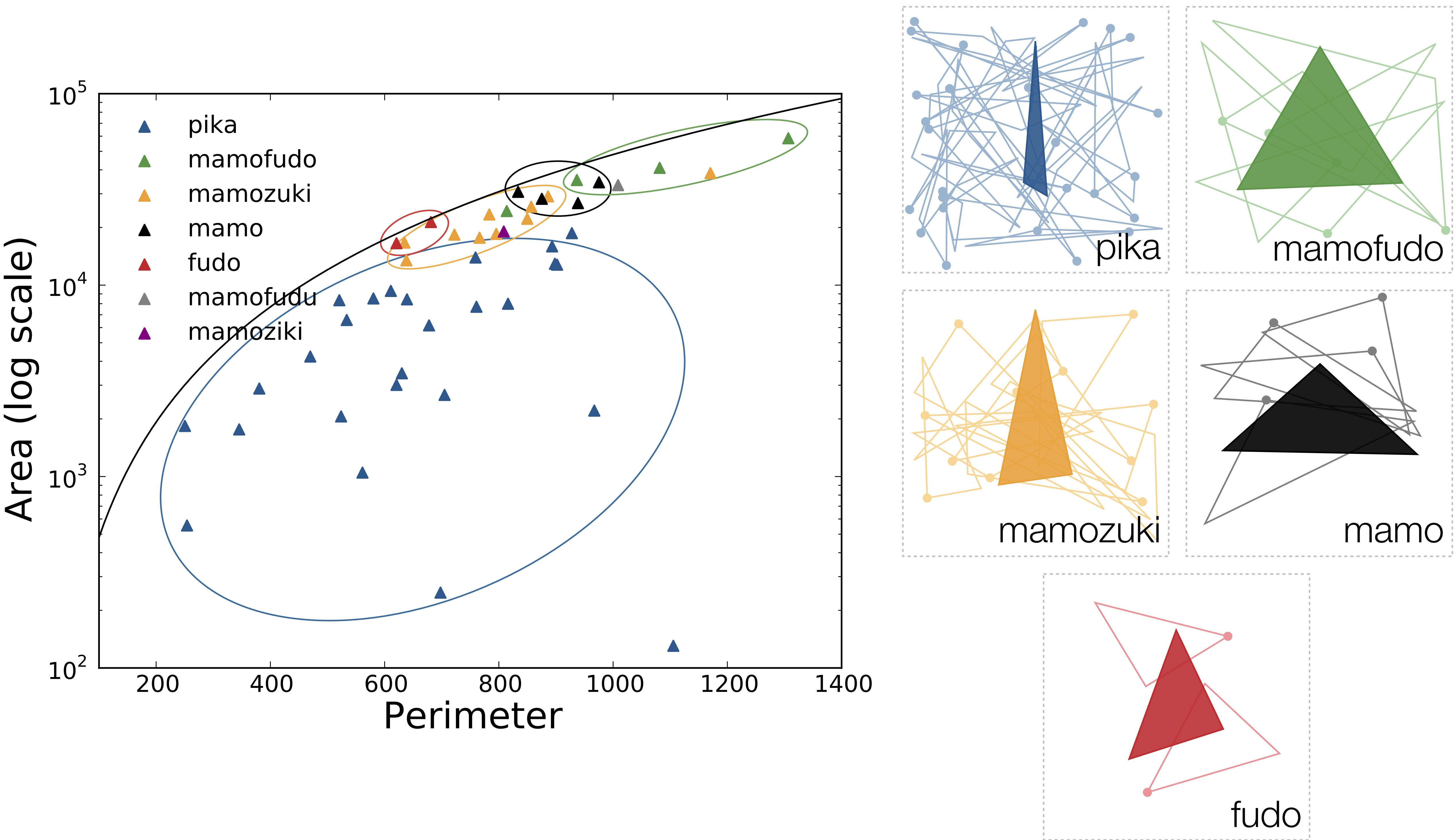



Chain A Chain B Chain C Chain D

# Results: Categorical structure



# Results: Categorical structure



# Results: Summary

**Hypothesis 1:** the languages will become increasingly learnable

$L = 1514$, $m = 4$, $n = 10$, $p < 0.001$ ✓

**Hypothesis 2:** categorical structure will emerge as a mechanism for circumventing the bottleneck on transmission

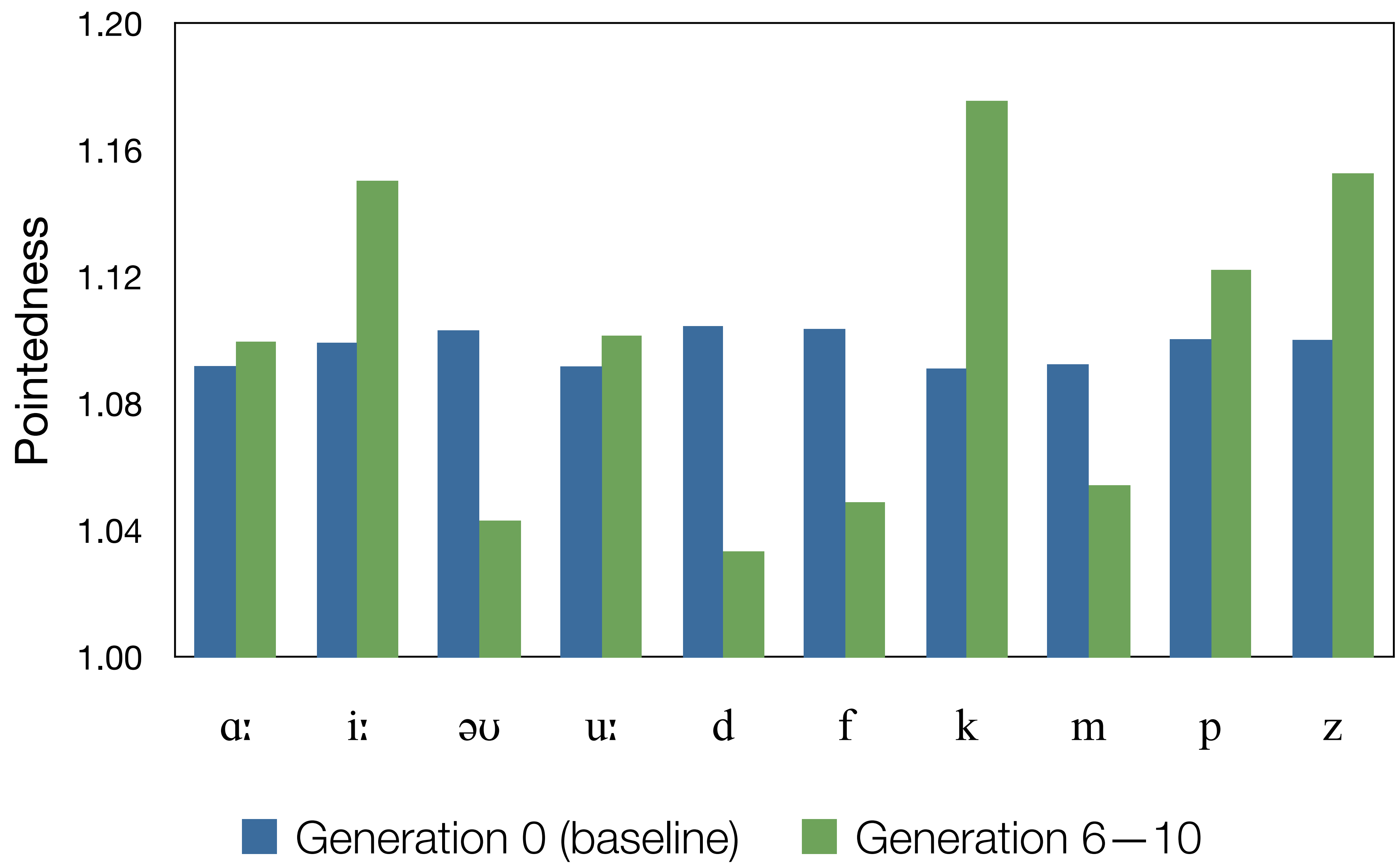
$L = 1461$, $m = 3$, $n = 11$, $p < 0.001$ ($d_T$ up to rigid motion)
$L = 1470$, $m = 3$, $n = 11$, $p < 0.001$ ($d_T$ up to scaled rigid motion) ✓

**Hypothesis 3:** an increase in learnability can be explained by an increase in structure

$r = 0.479$, $n = 36$, $p = 0.002$ ✓

# Results: Sound symbolism

Mean pointedness of triangles whose associated words contain phoneme *X*



# Summary

Experimental demonstration that categorical structure can arise from iterated learning

The meaning space has four key properties:

- **Continuous:** On each dimension, the triangle stimuli vary over a continuous scale
- **Vast in magnitude:** $6 \times 10^{15}$ possible triangle stimuli, vastly more than previous experiments
- **Complex dimensions:** Many possible dimensions to the space
- **Not pre-specified by the experimenter:** no particular hypothesis about which features participants would find salient

# Conclusions

Iterated learning in simple linear diffusion chains can give rise to categorical structure despite the fact that:

- stimuli never reoccur across participants
- there is no communicative pressure for expressivity

Although separate chains divided the space in subtly-different but lineage specific ways, participants showed a bias towards the shape and size properties

This suggests that iterated learning amplifies weak cognitive biases, giving rise to the categorical structure we observe in languages

Hannah Cornish

Simon Kirby

THE CARNEGIE TRUST FOR THE UNIVERSITIES OF SCOTLAND

BRITISH ACADEMY *for the humanities and social sciences*

# References


Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the USA*, *105*, 10681–10686.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, *10*, 707–710.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.

Matthews, C. (2009). The emergence of categorization: Language transmission in an iterated learning model using a continuous meaning space. (Unpublished master's dissertation). University of Edinburgh, Edinburgh, UK.

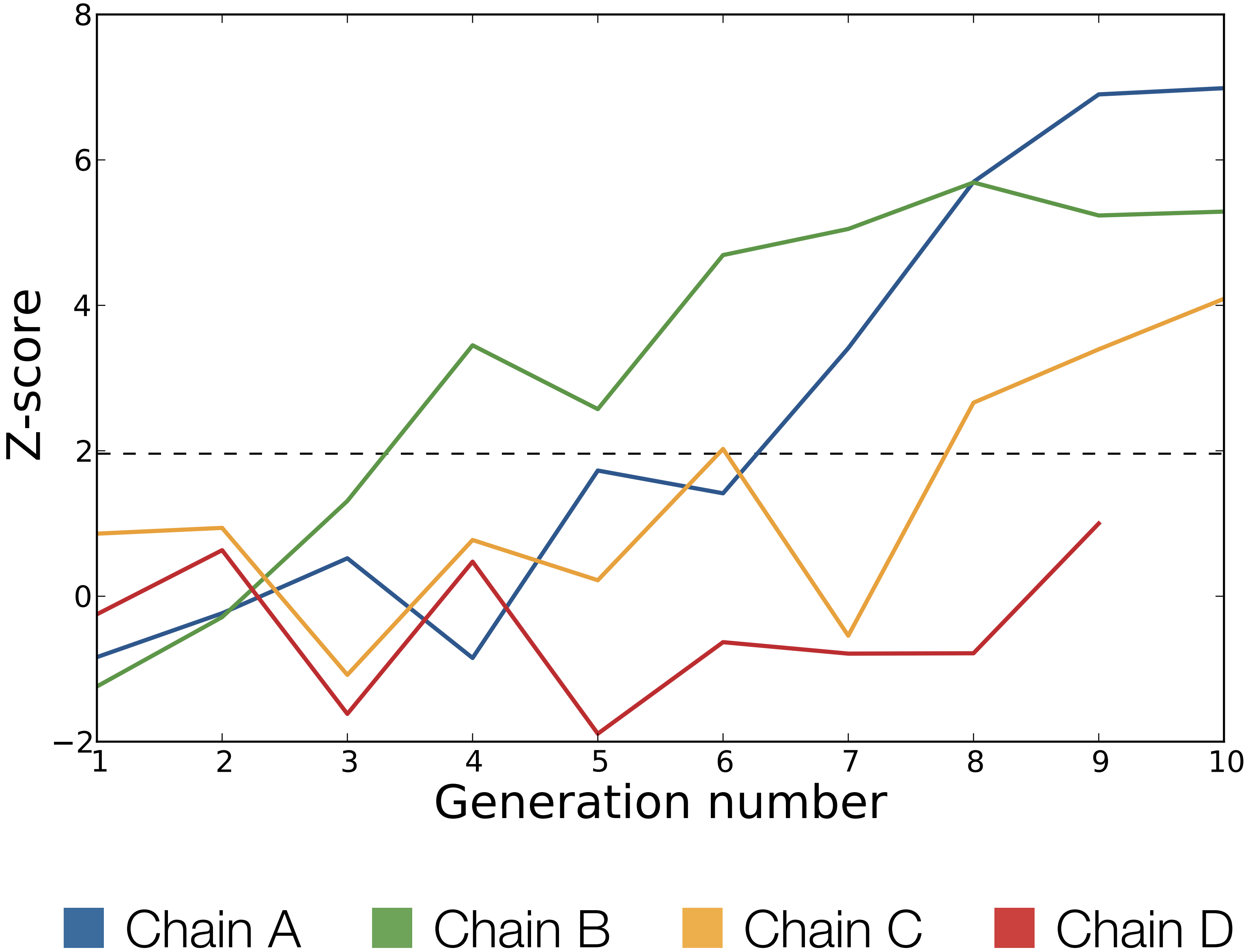
Page, E. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association*, *58*, 216–230.

Perfors, A., & Navarro, D. (2011). Language evolution is shaped by the structure of the world: An iterated learning analysis. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 477–482). Austin, TX: Cognitive Science Society.

Silvey, C., Kirby, S., & Smith, K. (2013). Communication leads to the emergence of sub-optimal category structures. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the Cognitive Science Society* (pp. 1312–1317). Austin, TX: Cognitive Science Society.

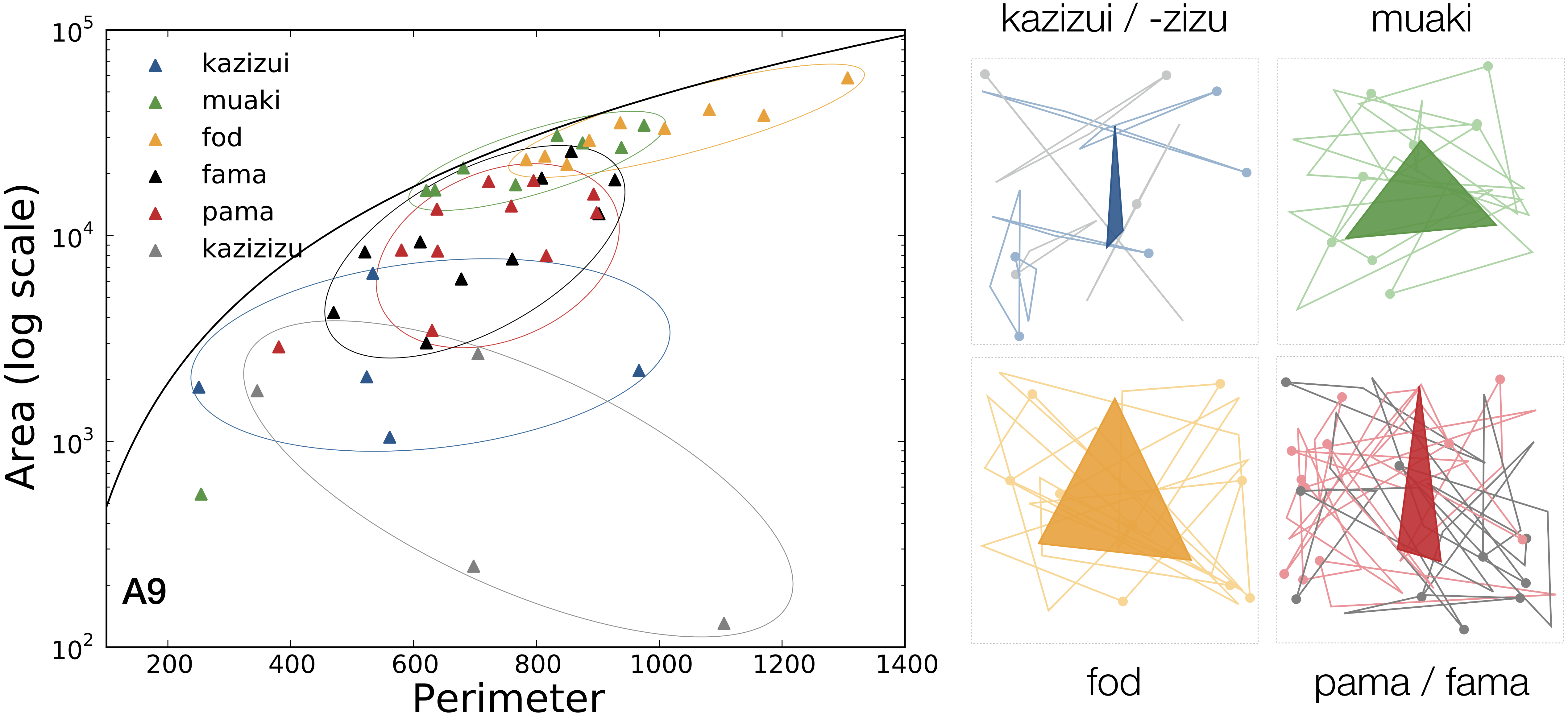
# Learnability

Transformation of transmission error scores to account for chance




$L = 1038$, $m = 4$, $n = 9$, $p < 0.001$

# Emergent language in chain A (gen 9)



# Emergent language in chain C (gen 8)